

PCTWORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁵ : C07H 15/12, C12N 1/00 C12P 21/06, C07K 3/00 A61K 37/22, C12Q 1/00	A1	(11) International Publication Number: WO 91/17170 (43) International Publication Date: 14 November 1991 (14.11.91)
(21) International Application Number: PCT/US91/02954 (22) International Filing Date: 3 May 1991 (03.05.91) (30) Priority data: 520,368 4 May 1990 (04.05.90) US (60) Parent Application or Grant (63) Related by Continuation US 520,368 (CIP) Filed on 4 May 1990 (04.05.90) (71) Applicant (for all designated States except US): BIOGEN, INC. [US/US]; 14 Cambridge Center, Cambridge, MA 02142 (US). (72) Inventors; and (75) Inventors/Applicants (for US only) : PEPINSKY, R., Blake [US/US]; 69 Parker Street, Apt. 2, Watertown, MA 02172 (US). ROSA, Margaret, D. [US/US]; 32 Grove Street, Winchester, MA 01890 (US). STOSSEL, Thomas, P. [US/US]; 12 Blake Street, Belmont, MA 02178 (US).		(74) Agents: HALEY, James, F., Jr. et al.; Fish & Neave, 875 Third Avenue, 29th Floor, New York, NY 10022 (US). (81) Designated States: AT (European patent), AU, BB, BE (European patent), BF (OAPI patent), BG, BJ (OAPI patent), BR, CA, CF (OAPI patent), CG (OAPI patent), CH (European patent), CI (OAPI patent), CM (OAPI patent), DE (European patent), DK (European patent), ES (European patent), FI, FR (European patent), GA (OAPI patent), GB (European patent), GR (European patent), HU, IT (European patent), JP, KP, KR, LK, LU (European patent), MC, MG, ML (OAPI patent), MR (OAPI patent), MW, NL (European patent), NO, PL, RO, SD, SE (European patent), SN (OAPI patent), SU, TD (OAPI patent), TG (OAPI patent), US. Published <i>With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>
(54) Title: MULTIMERIC GELSOLIN FUSION CONSTRUCTS (57) Abstract This invention relates to multimeric and hetero-multimeric gelsolin fusion constructs, compositions containing them and methods using them. More particularly, this invention relates to multimeric gelsolin fusion constructs in which at least two gelsolin fusion polypeptides are bound to vesicles containing polyphosphoinositides. This invention also relates to gelsolin fusion polypeptides which comprise gelsolin moieties linked to functional moieties and in particular, to CD4-gelsolin fusion polypeptides comprising an amino acid sequence for a human CD4 protein linked to a gelsolin moiety.		

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	ES	Spain	MG	Madagascar
AU	Australia	FI	Finland	ML	Mali
BB	Barbados	FR	France	MN	Mongolia
BE	Belgium	GA	Gabon	MR	Mauritania
BF	Burkina Faso	GB	United Kingdom	MW	Malawi
BG	Bulgaria	GN	Guinea	NL	Netherlands
BJ	Benin	GR	Greece	NO	Norway
BR	Brazil	HU	Hungary	PL	Poland
CA	Canada	IT	Italy	RO	Romania
CF	Central African Republic	JP	Japan	SD	Sudan
CG	Congo	KP	Democratic People's Republic of Korea	SE	Sweden
CH	Switzerland	KR	Republic of Korea	SN	Senegal
CI	Côte d'Ivoire	LI	Liechtenstein	SU	Soviet Union
CM	Cameroon	LK	Sri Lanka	TD	Chad
CS	Czechoslovakia	LU	Luxembourg	TG	Togo
DE	Germany	MC	Monaco	US	United States of America
DK	Denmark				

- 1 -

MULTIMERIC GELSOLIN FUSION CONSTRUCTSTECHNICAL FIELD OF INVENTION

This invention relates to multimeric and
5 hetero-multimeric gelsolin fusion constructs,
compositions containing them and methods using them.
More particularly, this invention relates to multimeric
gelsolin fusion constructs in which at least two
gelsolin fusion polypeptides are bound to vesicles
10 containing polyphosphoinositides. This invention also
relates to gelsolin fusion polypeptides which comprise
gelsolin moieties linked to functional moieties and, in
particular, to CD4-gelsolin fusion polypeptides
comprising an amino acid sequence for a human CD4
15 protein linked to a gelsolin moiety.

BACKGROUND ART

The rapid development of biotechnologies has
led to novel delivery and carrier systems for
pharmaceuticals, vaccines, diagnostics and other
20 bioactive molecules. Optimally, these systems enhance
the properties of the molecules they carry, complement
those molecules with characteristics they lack and
combine useful characteristics of different molecules.
Of particular interest to researchers are the serum
25 half-life of bioactive molecules, their affinity for
target particles and cells, targetability of bioactive

- 2 -

molecules, bioactivity, immunogenicity and the ability to administer or deliver several molecules simultaneously. Scientists are seeking to identify new molecules, including proteins, that they can
5 advantageously develop into these systems.

Gelsolin is a protein found in mammals and other vertebrates [H.L. Yin and T.P. Stossel, "Control of Cytoplasmic Actin Gel-sol Transformation by Gelsolin, a Calcium-dependent Regulatory Protein",
10 Nature, 281, pp. 583-86 (1979); F.S. Southwick and M.J. DiNubile, "Rabbit Alveolar Macrophages Contain a Ca^{2+} -sensitive, 41,000-dalton Protein Which Reversibly Blocks the 'Barbed' Ends of Actin Filaments but Does not Sever Them", J. Biol. Chem., 261, pp. 14191-95
15 (1986); T. Ankenbauer et al., "Proteins Regulating Actin Assembly in Oogenesis and Early Embryogenesis of Xenopus laevis: Gelsolin Is the Major Cytoplasmic Actin-binding Protein", J. Cell Biol., 107, pp. 1489-98 (1988); H.L. Yin et al., "Identification of Gelsolin, a
20 Ca^{2+} -dependent Regulatory Protein of Actin Gel-sol Transformation and Its Intracellular Distribution in a Variety of Cells and Tissues", J. Cell. Biol., 91, pp. 901-06 (1980); C.W. Dieffenbach et al., "Cloning of Murine Gelsolin and Its Regulation During
25 Differentiation", J. Biol. Chem., 264, pp. 13281-88 (1989)]. In mammals, gelsolin occurs in two forms -- a cytoplasmic form and a serum form. Gelsolin regulates the activity of actin, a major protein involved in cell structure and movement. Actin is a globular protein
30 with a slightly elongated shape that can polymerize into filaments. Polymerization occurs when the "barbed" end of one actin monomer binds non-covalently and reversibly to the "pointed" end of another. Inside most cells, monomers and short filaments exist in a
35 fluid-like "sol" state until the monomers are activated to polymerize into filaments and the filaments, in

- 3 -

turn, are activated to crosslink, producing a firmer "gel" phase that forms part of the cellular cytoskeleton. Investigators have observed that in the presence of calcium ion, gelsolin prevents the transition of monomers and filaments from gel phase to sol phase.

Gelsolin acts on actin in three ways. First, it severs the noncovalent bonds between the actin monomers that compose actin filaments ("severing"). Second, it binds to the barbed end of actin filaments and prevents elongation of the filament from that end ("capping"). Third, it binds to actin monomers and promotes the formation of actin filaments by providing a nucleus for polymerization ("nucleation"). The result is a steady state which favors short actin filaments unable to support the gel phase [P.A. Janmey et al., "Interactions of Gelsolin and Gelsolin-actin Complexes with Actin. Effects of Calcium on Actin Nucleation, Filament Severing, and End Blocking", Biochemistry, 24, pp. 3714-23 (1985)].

Gelsolin's actin-severing function is stoichiometric: one gelsolin molecule binds to two monomers on the actin filament, breaks the filament, and remains bound to both monomers. The binding of gelsolin to one of the monomers is Ca^{++} dependent, and chelating agents such as EGTA cause dissociation of gelsolin from only one monomer.

Scientists have identified two phosphatidyl inositol phosphate phospholipids that bind to and regulate the function of gelsolin. They are phosphatidylinositol 4-monophosphate (PIP) and phosphatidylinositol 4,5-biphosphate (PIP_2) [P.A. Janmey et al., "Polyphosphoinositide Micelles and Polyphosphoinositide-containing Vesicles Dissociate Endogenous Gelsolin-actin Complexes and Promote Actin Assembly from the Fast-growing End of Actin Filaments

- 4 -

Blocked by Gelsolin", J. Biol. Chem., 262, pp. 12228-36 (1987), P.A. Janmey and T.P. Stossel, "Modulation of Gelsolin Function by Phosphatidylinositol 4,5-biphosphate", Nature, 325, pp. 362-64 (1987) and P.A. Janmey and T.P. Stossel, "Gelsolin-phosphoinositide Interaction", J. Biol. Chem., 264, pp. 4825-31 (1989)]. These polyphosphoinositides are minor membrane phospholipids that play a role in signal transduction in cells [B. Alberts et al., Molecular Biology of the Cell, Second Edition, Garland Publishing, Inc., New York, New York, pp. 702-703 (1989)]. Together they comprise less than 10% of the total phospholipids of cell membranes, and PIP₂ comprises less than 1%. These two molecules inhibit gelsolin activity by binding to gelsolin and displacing the actin monomers that are bound to it in a non-Ca⁺⁺ dependent manner.

In extensively sonicated aqueous suspensions, both PIP and PIP₂ form vesicles. PIP₂ forms small vesicles, also called micelles, of about 80 nm in diameter, that contain about one-hundred PIP₂ molecules. Each PIP₂ micelle binds about eight gelsolin molecules. PIP forms larger unilamellar (one-layered) vesicles. Aggregation of PIP₂ into large unilamellar or multimellar vesicles in the presence of millimolar concentrations of Mg⁺⁺ or nonionic detergents decreases the ability of PIP₂ to inhibit the actin filament-severing function of gelsolin. Incorporation of PIP₂ into mixed vesicles composed of phosphatidylcholine (PC) also decreases this ability, although about a third of maximal activity persists, even in vesicles containing a very high ratio of PC to PIP₂. Mixed lipid vesicles whose composition approximates that of the cell membrane (less than 3% PIP₂) also inhibit gelsolin activity. Several other polyphosphoinositides which may be constructed, or have

- 5 -

already been identified in nature, would also be expected to bind gelsolin.

The cDNA for human plasma gelsolin encodes a protein of 755 amino acids plus a 27 amino acid signal sequence [Kwiatkowski et al., "Plasma and Cytoplasmic Gelsolins Are Encoded by a Single Gene and Contain a Duplicated Actin-binding Domain", Nature, 323, pp. 455-58 (1986)]. This cDNA sequence accounts for both the plasma and serum forms of gelsolin, which are the result of alternative transcriptional initiation sites and message processing from a single gene, 70 kb long [D. Kwiatkowski et al., "Genomic Organization and Biosynthesis of Secreted and Cytoplasmic Forms of Gelsolin", J. Cell Biol., 106, pp. 375-84 (1988)]. The difference between the plasma and cytoplasmic forms is a 25 amino-acid residue extension on plasma gelsolin. This appears to account for the difference in relative molecular weight between the proteins as assessed by SDS-polyacrylamide gel electrophoresis (SDS-PAGE), 93 kD and 90 kD, respectively.

Investigators have identified several functional domains of gelsolin [H.L. Yin et al., "Identification of a Polyphosphoinositide-modulated Domain in Gelsolin Which Binds to the Sides of Actin Filaments", J. Cell Biol., 106, pp. 805-12 (1988) and D. Kwiatkowski et al., "Identification of Critical Functional and Regulatory Domains in Gelsolin", J. Cell Biol., 108, pp. 1717-26 (1989)]. The gelsolin cDNA contains a strong tandem repeat that divides the molecule into two roughly equal halves. These structural halves correspond to two functional halves: The amino-terminal half of the protein contains a Ca^{++} -insensitive actin-severing function and the carboxy-terminal half has a Ca^{++} -sensitive actin binding domain. Within these two tandem repeats are six domains of weaker homology. The polypeptide has three actin

- 6 -

binding sites. Two monomer binding sites are located between residues 26-139 and 407-756 (probably 661-738) and an actin filament binding site is located between residues 151-406. Amino acid residues 732-738 are
5 potentially important for Ca^{++} regulation. Residues 660-738 are important for nucleation. This function probably requires actin binding sites on both halves of the molecule. The severing function resides in residues 1-160, possibly between residues 139-160, with
10 critical dependence on the sequence 150-160 (the first eleven residues of domain two). The PIP_2 -regulation of gelsolin's severing activity apparently resides within the first 160 residues. Sequences in domains 2 and 3 appear to hide a cryptic Ca^{++} -sensitive domain because
15 when they are removed, the severing function of gelsolin becomes Ca^{++} dependent.

Significantly, the amino acid sequence of gelsolin exhibits homology with several other actin binding proteins. It is forty-five percent homologous
20 with villin, found in vertebrate brush border microvilli, which also has a Ca^{++} -dependent actin severing function. It is thirty-three percent homologous with severin and fragmin [P. Matsudaira and P. Janmey, "Pieces in the Actin-severing Protein
25 Puzzle", Cell, 54, pp. 139-40 (1988)]. These polypeptides also bind PIP and PIP_2 .

Despite advances in biotechnology, the need still exists for methods and products which optimize the characteristics and delivery of pharmaceuticals,
30 vaccines, diagnostics and bioactive molecules -- including polyvalency, affinity for a single target particle, serum half-life, bioactivity and, in some cases, immunogenicity. Furthermore, systems in which the component parts may be easily varied would be

- 7 -

especially useful because they would allow one to test for species with optimal characteristics.

SUMMARY OF THE INVENTION

This invention solves these problems by
5 providing multimeric and hetero-multimeric gelsolin fusion constructs. A multimeric gelsolin fusion construct is a vesicle comprising at least one polyphosphoinositide, such as PIP or PIP₂, to which gelsolin fusion polypeptides are bound. Gelsolin
10 fusion polypeptides comprise gelsolin moieties linked to functional moieties which may be pharmaceutical agents, vaccine agents, diagnostic agents or other bioactive molecules. Hetero-multimeric gelsolin fusion constructs comprise at least two different functional
15 moieties or gelsolin moieties.

Gelsolin is a particularly attractive candidate for attachment to lipid vesicles because it binds specifically and with great affinity to polyphosphoinositides. Other proteins, related to
20 gelsolin, which also specifically bind polyphosphoinositides may also be employed. Some examples are villin, fragmin, severin, profilin, cofilin, Cap42(a), gCap39, CapZ and destrin. Lipocortin (annexin) and DNaseI are other molecules
25 that bind polyphosphoinositides. Proteins that specifically bind other lipids may also be used, as well as proteins that bind lipids non-specifically.

The fusion constructs of this invention advantageously utilize the ability of
30 polyphosphoinositide vesicles to bind multiple copies of gelsolin fusion polypeptides. Consequently, in contrast to monomeric molecules, the bioactive molecules linked to them as functional moieties are characterized by one or more of the following:
35 polyvalency, increased serum half-life, affinity for

- 8 -

target particles or cells, greater bioactivity or immunogenicity, and targetability.

The present invention also provides gelsolin fusion polypeptides. Gelsolin fusion polypeptides
5 comprise gelsolin moieties fused or chemically coupled to a functional moiety. In particular, this invention provides CD4-gelsolin fusion polypeptides.

The lipid composition of a vesicle may also be varied to permit the production of vesicles varying
10 in fluidity, size, the number of gelsolin molecules that will bind to it and the rate of degradation in the blood stream.

Depending upon the choice of functional moiety, multimeric and hetero-multimeric gelsolin
15 fusion constructs are characterized by many uses. Recognition molecules, such as those containing the antigen binding site of antibodies, viral receptors or cell receptors, are useful as functional moieties to target fusion proteins to particular antigens. When
20 targeted in this manner, multimeric gelsolin fusion constructs are useful to block the binding of viruses to cells that results in infection, or the binding of cells to other cells that, for example, characterizes pathologic inflammation. Due to the multivalency of
25 the fusion constructs of this invention, we believe that they possess greater affinity for the target than monovalent molecules. In one embodiment of this invention, the functional moiety is the receptor on human lymphocytes, CD4, which is the target of the HIV
30 virus -- the causative agent of AIDS and ARC.

When hetero-multimeric fusion constructs comprise gelsolin fusion polypeptides having combinations of recognition molecules and toxins, anti-retroviral agents or radionuclides, they are
35 useful as therapeutic agents which search out and destroy their target.

- 9 -

Multimeric gelsolin fusion constructs with recognition molecules are also useful for signal enhancement in diagnostic assays. As large, multimeric molecules, they present many binding sites for reporter
5 molecules, such as horseradish peroxidase-conjugated antibodies. Alternatively, they may take the form of hetero-multimeric constructs, possessing both recognition molecules and multiple reporter groups.

When the functional moiety is one or more
10 immunogen from one or more infectious agent, the fusion proteins of this invention are useful in vaccines.

Also, multimeric gelsolin fusion constructs may be employed as agents with increased bioactivity when the functional group is an enzyme, substrate, or
15 inhibitor.

This invention also provides multimeric gelsolin fusion constructs that are liposomes whose constituents include polyphosphoinositides and that contain bioactive agents in their interiors.

20 This invention further provides DNA sequences that encode gelsolin fusion polypeptides, recombinant DNA molecules comprising them and unicellular host cells transformed with them. And this invention provides methods for producing these fusion
25 polypeptides by culturing such hosts.

This invention also provides compositions comprising any of the above-identified fusion polypeptides or proteins that are useful as therapeutic, prophylactic or diagnostic agents.
30 Multimeric CD4-gelsolin fusion constructs may be used in diagnosing, preventing and treating AIDS, ARC or HIV infection.

- 10 -

BRIEF DESCRIPTION OF THE DRAWINGS

Figures 1A-1F ("Figure 1") (SEQ ID NO:1) depict the DNA sequence and deduced amino acid sequence of human gelsolin as set forth in D.J. Kwiatkowski et al., Nature, 323, pp. 455-58 (1986). The negatively numbered amino acids correspond to the signal sequence, which is absent from the mature polypeptide. Throughout this specification, references to human gelsolin by amino acid sequence or DNA sequence correspond to the coordinate system set forth in this figure.

Figure 2 depicts the functional regions of human gelsolin amino acid sequence.

Figures 3A-3D ("Figure 3") (SEQ ID NO:2) depict the DNA sequence and deduced amino acid sequence of human CD4 DNA. Nucleotides 1-75 are derived from plasmid p170.2. Nucleotides 76-741 are derived from plasmid pCD4-gelsolin. Nucleotides 742 to 1377 are derived from p170.2. Throughout this specification, references to CD4 by amino acid or DNA sequence correspond to the coordinate system of this figure, unless otherwise specified.

Figure 4 depicts the domain structure of human CD4 protein. The numbered amino acids are cysteine residues involved in disulfide bonding according to Figure 3.

Figure 5 depicts the DNA sequences of the oligomers used in the processes set forth in the examples of this application. The gelsolin sequences in this figure are derived from SEQ ID NO:1. ACE 144 is SEQ ID NO:3. ACE 145 is SEQ ID NO:4. T4 AID-133 is SEQ ID NO:5. T4AID-134 is SEQ ID NO:6. T4AID-137 is SEQ ID NO:7. T4AID-176 is SEQ ID NO:8. T4AID-176 is SEQ ID NO:9.

Figure 6 depicts the construction of plasmid pCD4-gelsolin.

- 11 -

Figures 7A-7B ("Figure 7") (SEQ ID NO:10) depicts the DNA sequence and deduced amino acid sequence of pCD4-gelsolin.

Figure 8 is a restriction map of
5 pCD4-gelsolin.

Figure 9 depicts the construction of plasmid pDC219.

Figures 10A-10F ("Figure 10") (SEQ ID NO:11) depict the DNA sequence of p218-8.

10 Figure 11 depicts the construction of plasmid p λ P_L180cys.

Figures 12A-12I ("Figure 12") (SEQ ID NO:12) depict the DNA sequence of pBG391.

Figures 13A-13H ("Figure 13") (SEQ ID NO:13)
15 depict the DNA sequence of pEX46.

DETAILED DESCRIPTION OF THE INVENTION

"Human plasma gelsolin" refers to a polypeptide having the amino acid sequence depicted in Figure 1 (SEQ ID NO:1) from amino acids -27 to +755.
20 It should be understood that polypeptide expression often involves post-translational modifications such as cleavage of the signal sequence, intramolecular disulfide bonding, glycosylation and the like. The use of the term, human plasma gelsolin, contemplates such
25 modifications to the amino acid sequence of Figure 1 (SEQ ID NO:1). The term also includes gelsolin obtained from natural, recombinant or synthetic sources.

"Multimeric gelsolin fusion constructs" and
30 "hetero-multimeric gelsolin fusion constructs" each comprise gelsolin fusion polypeptides bound to a vesicle of aggregated phospholipids. A "gelsolin fusion polypeptide" comprises a gelsolin moiety bound to a functional moiety. "Functional moieties" may be
35 polypeptides ("polypeptide moieties") or chemical

- 12 -

compounds ("chemical moieties"). Throughout this application, specific gelsolin fusion polypeptides are referred to by the name of the functional moiety. For example, we call a gelsolin fusion polypeptide having
5 CD4 as the functional moiety, CD4-gelsolin fusion polypeptide. Hetero-multimeric gelsolin fusion constructs comprise at least two different functional moieties or gelsolin moieties.

When the functional moiety is a polypeptide,
10 gelsolin fusion polypeptides may be produced by chemical crosslinking or genetic fusion. Genetic fusion involves creating a hybrid DNA sequence in which the DNA sequence encoding the polypeptide is fused to the 5' end or 3' end of a DNA sequence encoding the
15 gelsolin moiety. Upon expression in an appropriate host, this hybrid DNA sequence produces a gelsolin fusion polypeptide in which the polypeptide moiety is fused to the N-terminus or C-terminus of the gelsolin moiety.

20 A "gelsolin moiety" as used herein is gelsolin or a fragment thereof that specifically binds to a polyphosphoinositide. Preferably, the gelsolin moiety will be derived from human plasma gelsolin. A gelsolin moiety preferably includes amino acids +150 to
25 +160 of Figure 1 (SEQ ID NO:1). As demonstrated herein, the polypeptide containing amino acids +150 to +169 of Figure 1 (SEQ ID NO:1) has the ability to bind PIP_2 . We believe that gelsolin derived from non-human vertebrates may also be useful according to this
30 invention. The structure of gelsolin is highly conserved in evolution and gelsolin from non-human mammals may not be immunogenic in humans.

Lipid binding proteins ("LBPs") other than gelsolin are also known to exist. These proteins, or
35 fragments of them that bind to particular lipids, are useful as LBP moieties (similarly to gelsolin moieties)

- 13 -

- to produce LBP fusion polypeptides that bind to vesicles containing the particular lipid. This creates multimeric or hetero-multimeric LBP fusion constructs. Gelsolin-related proteins that specifically bind
- 5 polyphosphoinositides include villin, severin, fragmin, profilin, cofilin, Cap42(a), gCap39, CapZ and destrin [E. Andr   et al., "Severin, Gelsolin, and Villin Share a Homologous Sequence in Regions Presumed to Contain F-actin Severing Domains", J. Biol. Chem., 263,
10 pp. 722-27 (1988); W.L. Bazari et al., "Villin Sequence and Peptide Map Identify Six Homologous Domains", Proc. Natl. Acad. Sci., USA, 85, pp. 4986-90 (1988); C. Ampe et al., "The Primary Structure of Human Platelet Profilin: Reinvestigation of the Calf Spleen Sequence",
15 FEBS Letters, 228, pp. 17-21 (1988); D.J. Kwiatkowski and G.A.P. Bruns, "Human Profilin", J. Biol. Chem., 263, pp. 5910-15 (1988); I. Lassing and U. Lindberg, "Specificity of the Interaction Between Phosphatidylinositol 4,5-biphosphate and the Profilin:
20 Actin Complex", J. Cell. Biochem., 37, pp. 255-67 (1988); C. Ampe and J. Vandekerckhove, "The F-actin Capping Proteins of Physarum polycephalum", EMBO. J., 6, pp. 4149-57 (1987); I. Lassing and U. Lindberg, "Specific Interaction between Phosphatidylinositol 4,5-
25 biphosphate and Profilactin", Nature, 314, pp. 472-74 (1985), F.-X. Yu et al., "gCap39, a Calcium Ion- and Polyphosphoinositide-regulated Actin Capping Protein", Science, 250, pp. 1413-15 (1990); and N. Yonezawa et al., "Inhibition of the Interactions of Cofilin,
30 Destrin and Deoxyribonuclease I with Actin by Phosphoinositides", J. Biol. Chem., 265, pp. 8382-86 (1990)]. Other LBPs that specifically bind polyphosphoinositides are lipocortin [K. Machoczek et al., "Lipocortin I and Lipocortin II Inhibit
35 Phosphoinositide and Polyphosphoinositide-specific Phospholipase C" FEBS Letters, 251, pp. 207-12 (1989)]

- 14 -

and DNase I [J.A. Cooper et al., "The Role of Actin Polymerization in Cell Motility", Ann. Rev. Phys., 53, pp. 585-605 (1991)]. Protein kinase C is also an LBP which binds to some phospholipids.

5 DNA sequences encoding gelsolin moieties are derived from DNA sequences encoding gelsolin. Several methods are available to obtain these DNA sequences. First, one can chemically synthesize the gelsolin gene or a degenerate version of it using a commercially
10 available chemical synthesizer. Figure 1 (SEQ ID NO:1) sets forth a DNA sequence for gelsolin. The coding region encompasses nucleotides +1 to +2360.

Second, one can isolate a cDNA sequence encoding gelsolin by screening a cDNA library. Many
15 screening methods are known to the art. For example, colonies may be screened by nucleic acid hybridization with oligonucleotide probes. Probes may be prepared by chemically synthesizing an oligonucleotide having part of the known DNA sequence of gelsolin. Alternatively,
20 cDNA libraries may be constructed in expression vectors, such as λ gt11, and the colonies screened with anti-gelsolin antibodies.

Third, one can isolate a cDNA encoding gelsolin or a gelsolin moiety by amplifying DNA with
25 polymerase chain reaction (PCR). We describe this process in Example I.

The DNA sequence encoding the gelsolin moiety may then be fused to a DNA sequence encoding the polypeptide moiety. DNA sequences for the polypeptide
30 moieties useful in this invention are available from many sources. These include DNA sequences described in the literature and DNA sequences for particular polypeptides obtained by any of the conventional molecular cloning techniques.

35 A wide array of polypeptides are useful to produce the gelsolin fusion polypeptides of this

- 15 -

invention. Those most useful include polypeptides that are advantageously administered in multimeric form. For example, viral receptors, cell receptors or cell ligands are useful because they typically bind to particles or cells exhibiting many copies of the receptor. Fusion constructs containing these fusion polypeptides are useful in therapies that involve the inhibition of viral-cell or cell-cell binding. Useful viral-cell receptors include ICAM1, a rhinovirus receptor; the polio virus receptor [J. M. White and D.R. Littman, "Viral Receptors of the Immunoglobulin Superfamily", Cell, 56, pp. 725-28 (1989)] and, most preferably, CD4, the HIV receptor. Cell-cell receptors or ligands include members of the vascular cell adhesion molecule family, such as ICAM1, ELAM1, VCAM1 and VCAM1b and their lymphocyte counterparts (ligands) LFA1, CDX and VLA4. These molecules are involved in pathologic inflammation [M.P. Bevilacqua et al., "Identification of an Inducible Endothelial-Leukocyte Adhesion Molecule", Proc. Natl. Acad. Sci., USA, 84, pp. 9238-42 (1987); L. Osborn et al., "Direct Expression Cloning of Vascular Cell Adhesion Molecule 1: A Cytokine-induced Endothelial Protein that Binds to Lymphocytes", Cell, 59, pp. 1203-11 (1989); C.A. Hession et al., "Endothelial Cell-leukocyte Adhesion Molecules (ELAMs) and Molecules Involved in Leukocyte Adhesion (MILAs)", WO 90/13300]. Other lymphocyte associated antigens, such as LFA2 (CD2) and LFA3 (both members of the CD11/CD18 family) and PAgEM are also useful.

Bacterial immunogens, parasitic immunogens and viral immunogens may be used as polypeptide moieties to produce multimeric or hetero-multimeric gelsolin fusion constructs useful as vaccines. Bacterial sources of these immunogens include those responsible for bacterial pneumonia and pneumocystis

- 16 -

pneumonia. Parasitic sources include malarial parasites, such as Plasmodium. Viral sources include poxviruses, e.g., cowpox virus and orf virus; herpes viruses, e.g., herpes simplex virus type 1 and 2, 5 B-virus, varicella-zoster virus, cytomegalovirus, and Epstein-Barr virus; adenoviruses, e.g., mastadenovirus; papovaviruses, e.g., papillomaviruses, and polyomaviruses such as BK and JC virus; parvoviruses, e.g., adeno-associated virus; reoviruses, e.g., 10 reoviruses 1, 2 and 3; orbiviruses, e.g., Colorado tick fever; rotaviruses, e.g., human rotaviruses; alphaviruses, e.g., Eastern encephalitis virus and Venezuelan encephalitis virus; rubiviruses, e.g., rubella; flaviviruses, e.g., yellow fever virus, Dengue 15 fever viruses, Japanese encephalitis virus, Tick-borne encephalitis virus and hepatitis C virus; coronaviruses, e.g., human coronaviruses; paramyxoviruses, e.g., parainfluenza 1, 2, 3 and 4 and mumps; morbilliviruses, e.g., measles virus; 20 pneumovirus, e.g., respiratory syncytial virus; vesiculoviruses, e.g., vesicular stomatitis virus; lyssaviruses, e.g., rabies virus; orthomyxoviruses, e.g., influenza A and B; bunyaviruses e.g., LaCrosse virus; phleborviruses, e.g., Rift Valley fever virus; 25 nairoviruses, e.g., Congo hemorrhagic fever virus; hepadnaviridae, e.g., hepatitis B; arenaviruses, e.g., lcm virus, Lassa virus and Junin virus; retroviruses, e.g., HTLV I, HTLV II, HIV I and HIV II; enteroviruses, e.g., polio virus 1, 2 and 3, coxsackie viruses, 30 echoviruses, human enteroviruses, hepatitis A virus, hepatitis E virus, and Norwalk virus; rhinoviruses e.g., human rhinovirus; and filoviridae, e.g., Marburg (disease) virus and Ebola virus.

Immunoglobulins or fragment thereof that bind 35 to a target molecule may also be employed as functional moieties. Immunoglobulin molecules are bivalent, but

- 17 -

multimeric immunoglobulin-gelsolin fusion constructs, which are multivalent, may demonstrate increased affinity or avidity for the target. Investigators have also made use of single domain antibodies (dAbs) [E.S. Ward et al., "Binding Activities of a Repertoire of Single Immunoglobulin Variable Domains Secreted from Escherichia coli", Nature, 341, pp. 544-46 (1989)]. One can generate monoclonal Fab fragments recognizing specific antigens using the technique of Huse et al. and use individual domains as functional moieties in multimeric or hetero-multimeric gelsolin fusion constructs according to this invention [W.D. Huse et al., "Generation of a Large Combinatorial Library of the Immunoglobulin Repertoire in Phage Lambda", Science, 246, pp. 1275-81 (1989)]. See also A. Skerra and A. Pluckthun "Assembly of a Functional Immunoglobulin Fv Fragment in Escherichia coli", Science, 240, pp. 1038-43 (1988)].

According to this invention, multimeric gelsolin fusion constructs may be produced in which the functional moiety is an enzyme, enzyme substrate or enzyme inhibitor. We believe that such agents will exhibit greater bioactivity than monomeric molecules because multimers have a higher density of the moiety and will exhibit increased catalytic rate. For example, we believe that a multimeric gelsolin fusion construct with tissue plasminogen activator would have greater clot-dissolving catalytic activity than its monovalent counterpart. Similarly, we believe that a multimeric gelsolin fusion construct with hirudin would demonstrate greater anti-coagulant activity than hirudin alone.

Other useful functional moieties include, but are not limited to, polypeptides such as cytokines, including the various IFN- α 's, particularly $\alpha 2$, $\alpha 5$, $\alpha 8$, IFN- β and IFN- γ , the various interleukins, including

- 18 -

IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7 and IL-8 and the tumor necrosis factors, TNF- α , and β . In addition, functional moieties include monocyte colony stimulating factor (M-CSF), granulocyte colony stimulating factor (G-CSF), granulocyte macrophage colony stimulating factor (GM-CSF), erythropoietin, platelet-derived growth factor (PDGF), and human and animal hormones, including growth hormones and insulin.

According to one embodiment of this invention, multimeric gelsolin fusion constructs comprise CD4-gelsolin fusion polypeptides. CD4 is the receptor on those white blood cells, T-lymphocytes, which recognizes HIV, the causative agent of AIDS and ARC [P.J. Maddon et al., "The T4 Gene Encodes the AIDS Virus Receptor and Is Expressed in the Immune System and the Brain", Cell, 47, pp. 333-48 (1986)]. Specifically, CD4 recognizes the HIV viral surface protein, gp120 and gp160. In CD4-gelsolin fusion polypeptides the functional moiety is a polypeptide moiety comprising full length CD4 or a fragment thereof, preferably soluble CD4. Use of the term, CD4, in this specification may refer to full length CD4 or fragments of CD4, unless specified.

A DNA sequence encoding full length human CD4 polypeptide and its deduced amino acid sequence is set forth in Figure 3 (SEQ ID NO:2). (See also P.J. Maddon et al., "The Isolation and Nucleotide Sequence of a cDNA Encoding the T Cell Surface Protein T4: A New Member of the Immunoglobulin Gene Family", Cell, 42, pp. 93-104 (1985).) Based upon its deduced primary structure, the CD4 polypeptide is divided into functional domains as follows:

- 19 -

	<u>Structure/Proposed Location</u>	<u>Amino Acid Coordinates In Figure 3</u>
	Hydrophobic/Secretory Signal	-25 to -1
5	First Immunoglobulin-related domain/Extracellular	+1 to +107
	Second Immunoglobulin-related domain/Extracellular	+108 to +177
10	Third Immunoglobulin-related domain/Extracellular	+178 to +293
	Fourth Immunoglobulin-related domain/Extracellular	+294 to +370
	Hydrophobic/Transmembrane Sequence	+371 to +391
15	Very Hydrophilic/ Intracytoplasmic	+392 to +431

The first immunoglobulin-related domain can be further resolved into a variable-related (V) region and joint-related (J) region, beginning at about amino acid +95

20 [S.J. Clark et al., "Peptide and Nucleotide Sequences of Rat CD4 (W3/25) Antigen: Evidence for Derivation from a Structure with Four Immunoglobulin-related Domains", Proc. Natl. Acad. Sci., USA, 84, pp. 1649-53 (1987)].

25 These domains also correspond roughly to structural domains of the CD4 protein due to intra-domain disulfide bonding. Thus, disulfide bonds join amino acids at positions +16 and +84 in the first immunoglobulin-related domain, amino acids +130 and

30 +159 of the second immunoglobulin-related domain, and amino acids +303 and +345 of the fourth immunoglobulin-related domain. Figure 4 depicts the domain structure of the full length human CD4 protein.

35 Soluble CD4 proteins have been constructed by truncating the full length CD4 protein at amino acid +375, to eliminate the transmembrane and cytoplasmic

- 20 -

domains. Such proteins have been produced by recombinant DNA techniques and are referred to as recombinant soluble CD4 (rsCD4) [R.A. Fisher et al., "HIV Infection Is Blocked In Vitro by Recombinant Soluble CD4", Nature, 331, pp. 76-78 (1988); Fisher et al., PCT patent application WO 89/01940 (incorporated herein by reference)]. These soluble CD4 proteins advantageously interfere with the CD4⁺ lymphocyte/HIV interaction by blocking or competitive binding mechanisms which inhibit HIV infection of cells expressing the CD4 protein. The first immunoglobulin-related domain is sufficient to bind gp120 and gp160. By acting as soluble virus receptors, soluble CD4 proteins are useful as antiviral therapeutics to inhibit HIV binding to CD4⁺ lymphocytes and virally induced syncytia formation.

The CD4 polypeptides useful in this invention include all CD4 polypeptides which bind to or otherwise inhibit gp120 and gp160. These include fragments of CD4 lacking the transmembrane domain, amino acids +371 to +391 of Figure 3 (SEQ ID NO:2). Such fragments may be truncated forms of CD4 or be fusion proteins in which the fourth immunoglobulin-related domain is joined directly to the hydrophilic cytoplasmic domain. We shall refer herein to a CD4 polypeptide which includes amino acids +1 to +X of Figure 3 (SEQ ID NO:2), and optionally including an N-terminal methionine or f-methionine, as "CD4(X)". When a CD4 polypeptide is engineered to include a carboxy-terminal cysteine, we shall refer to the polypeptide as "CD4(XCys)".

For example, referring now to Figure 3 (SEQ ID NO:2), a soluble CD4 protein containing the first immunoglobulin-like domain preferably will contain at least amino acids +1 to +84 and at most amino acids +1 to +129. Most preferably, a soluble CD4 protein

- 21 -

comprises amino acids +1 to +111 [CD4(111)]. A soluble CD4 protein containing the first two immunoglobulin-like domains preferably will include at least amino acids +1 to +159 and at most amino acids +1 to +302.

- 5 More preferably, a soluble CD4 protein will include at least amino acids +1 to +175 and at most amino acids +1 to +190. Most preferably, a soluble CD4 protein will include amino acids +1 to +181 [CD4(181)], amino acids +1 to +183 [CD4(183)], or amino acids +1 to +187
- 10 [CD4(187)]. A soluble CD4 protein which includes the first four immunoglobulin-like domains preferably will include at least amino acids +1 to +345 [CD4(345)] and at most amino acids +1 to +375 [CD4(375)]. Any of these molecules may optionally include the CD4 signal
- 15 sequence, amino acids -23 to -1 of Figure 3 (SEQ ID NO:2). Also, these molecules may have a modified methionine residue preceding amino acid, +1.

- Soluble CD4 proteins useful in the fusion polypeptides and methods of this invention may be
- 20 produced in a variety of ways. According to the coordinate system in Figure 3 (SEQ ID NO:2), the amino terminal amino acid of mature CD4 protein isolated from T cells is lysine, encoded at nucleotides 136 to 139 of Figure 3 (SEQ ID NO:2). [D.R. Littman et al.,
- 25 "Corrected CD4 Sequence", Cell, 55, p. 541 (1988).] Soluble CD4 proteins also include those in which amino acid +1 is asparagine, +62 is arginine and +229 is phenylalanine. Therefore, when we refer to CD4, we intend to include amino acid sequences including any or
- 30 all of these substitutions. Soluble CD4 polypeptides may be produced by conventional recombinant techniques involving oligonucleotide-directed mutagenesis and restriction digestion, followed by insertion of linkers, or by digesting full-length CD4 protein with
- 35 enzymes.

- 22 -

Soluble CD4 proteins include those produced by recombinant techniques according to the processes set forth in copending, commonly assigned United States patent applications Serial No. 094,322, filed
5 September 4, 1987 and Serial No. 141,649, filed January 7, 1988, and PCT patent application Serial No. PCT/US88/02940, filed September 1, 1988, and published as PCT patent application WO 89/01940, the disclosures of which are hereby incorporated by
10 reference.

Microorganisms and recombinant DNA molecules characterized by DNA sequences coding for soluble CD4 proteins are exemplified by cultures described in PCT patent application WO 89/01940. They were deposited in
15 the In Vitro International, Inc. culture collection, in Linthicum, Maryland, USA on September 2, 1987 and identified as:

EC100: E.coli JM83/pEC100 - IVI 10146
BG377: E.coli MC1061/pBG377 - IVI 10147
20 BG380: E.coli MC1061/pBG380 - IVI 10148
BG381: E.coli MC1061/pBG381 - IVI 10149.

Such microorganisms and recombinant DNA molecules are also exemplified by cultures deposited in the In Vitro International, Inc. culture collection on January 6,
25 1988 and identified as:

BG-391: E.coli MC1061/pBG391 - IVI 10151
BG-392: E.coli MC1061/pBG392 - IVI 10152
BG-393: E.coli MC1061/pBG393 - IVI 10153
BG-394: E.coli MC1061/pBG394 - IVI 10154
30 BG-396: E.coli MC1061/pBG396 - IVI 10155
203-5 : E.coli SG936/p203-5 - IVI 10156.

Additionally, such microorganisms and recombinant DNA molecules are exemplified by cultures deposited in the In Vitro International, Inc. culture
35 collection on August 24, 1988 and identified as:

- 23 -

211-11: E.coli A89/pBG211-11 - IVI 10183
214-10: E.coli A89/pBG214-10 - IVI 10184
215-7 : E.coli A89/pBG215-7 - IVI 10185.

Multimeric CD4-gelsolin fusion constructs
5 comprising CD4-gelsolin fusion polypeptides may be used
in pharmaceutical compositions and methods to treat
humans having AIDS, ARC, HIV infection, or antibodies
to HIV. Accordingly, they may be used to lessen the
immuno-compromising effects of HIV infection or to
10 prevent the incidence and spread of HIV infection. In
addition, these fusion proteins and methods may be used
for treating AIDS-like diseases caused by retroviruses,
such as simian immunodeficiency viruses, in mammals,
including humans.

15 DNA sequences encoding gelsolin fusion
polypeptides are useful for producing multimeric
gelsolin fusion constructs. The preferred process for
using these DNA sequences involves expressing the
gelsolin fusion polypeptide in an appropriate host,
20 isolating the polypeptide, and binding it to a vesicle
comprising a polyphosphoinositide.

As is well known in the art, for expression
of the DNA sequences of this invention, the DNA
sequence should be operatively linked to an expression
25 control sequence in an appropriate expression vector
and employed in that expression vector to transform an
appropriate unicellular host. Such operative linking
of a DNA sequence of this invention to an expression
control sequence, of course, includes the provision of
30 a translational start signal in the correct reading
frame upstream of the DNA sequence. If a particular
DNA sequence being expressed does not begin with an
ATG, the start signal will result in an additional
amino acid -- methionine (or f-methionine in
35 bacteria) -- being located at the N-terminus of the

- 24 -

product. While such methionyl-containing product may be employed directly in the compositions and methods of this invention, it is usually more desirable to remove the methionine before use. Methods are known to those
5 of skill in the art to remove such N-terminal methionines from polypeptides expressed with them. For example, certain hosts and fermentation conditions permit removal of substantially all of the N-terminal methionine in vivo. Expression in other hosts requires
10 in vitro removal of the N-terminal methionine. However, such in vivo and in vitro methods are well known in the art.

A wide variety of host/expression vector combinations may be employed in expressing the DNA
15 sequences of this invention. Useful expression vectors, for example, may consist of segments of chromosomal, non-chromosomal and synthetic DNA sequences, such as various known derivatives of SV40 and known bacterial plasmids, e.g., plasmids from
20 E.coli including colE1, pCR1, pBR322, pMB9 and their derivatives, wider host range plasmids, e.g., RP4, phage DNAs, e.g., the numerous derivatives of phage λ , e.g., NM989, and other DNA phages, e.g., M13 and filamentous single stranded DNA phages, yeast plasmids,
25 such as the 2 μ plasmid or derivatives thereof, and vectors derived from combinations of plasmids and phage DNAs, such as plasmids which have been modified to employ phage DNA or other expression control sequences.

In addition, any of a wide variety of
30 expression control sequences -- sequences that control the expression of a DNA sequence when operatively linked to it -- may be used in these vectors to express the DNA sequences of this invention. Such useful expression control sequences, include, for example, the
35 early and late promoters of SV40 or adenovirus, the lac system, the trp system, the TAC or TRC system, the

- 25 -

major operator and promoter regions of phage λ , the control regions of fd coat protein, the promoter for 3-phosphoglycerate kinase or other glycolytic enzymes, the promoters of acid phosphatase, e.g., Pho5, the
5 promoters of the yeast α -mating factors, the polyhedron promoter of the baculovirus system and other sequences known to control the expression of genes of prokaryotic or eukaryotic cells or their viruses, and various combinations thereof.

10 A wide variety of unicellular host cells are also useful in expressing the DNA sequences of this invention. These hosts include well known eukaryotic and prokaryotic hosts, such as strains of E.coli,
15 Pseudomonas, Bacillus, Streptomyces, fungi, such as yeasts, and animal cells, such as CHO and mouse cells, African green monkey cells, such as COS-1, COS-7, BSC 1, BSC 40, and BMT 10, insect cells, and human cells and plant cells in tissue culture. For animal cell expression, we prefer CHO cells and COS-7 cells.

20 It should of course be understood that not all vectors and expression control sequences will function equally well to express the DNA sequences of this invention. Neither will all hosts function
25 equally well with the same expression system. However, one of skill in the art may make a selection among these vectors, expression control sequences, and hosts without undue experimentation and without departing from the scope of this invention. For example, in
30 selecting a vector, the host must be considered because the vector must replicate in it. The vector's copy number, the ability to control that copy number, and the expression of any other proteins encoded by the vector, such as antibiotic markers, should also be considered.

35 In selecting an expression control sequence, a variety of factors should also be considered. These

- 26 -

include, for example, the relative strength of the system, its controllability, and its compatibility with the particular DNA sequence of this invention, particularly as regards potential secondary structures.

5 Unicellular hosts should be selected by consideration of their compatibility with the chosen vector, the toxicity of the product coded for on expression by the DNA sequences of this invention to them, their secretion characteristics, their ability to fold
10 proteins correctly, their fermentation requirements, and the ease of purification of the products coded on expression by the DNA sequences of this invention.

Within these parameters, one of skill in the art may select various vector/expression control
15 system/host combinations that will express the DNA sequences of this invention on fermentation or in large scale animal culture, e.g., CHO cells or COS-7 cells.

According to one embodiment of this invention, a plasmid comprising a DNA sequence encoding
20 a CD4-gelsolin fusion polypeptide operatively linked to a λP_L promoter expression control sequence is expressed in E.coli to produce a CD4-gelsolin fusion polypeptide.

The polypeptides and proteins produced on expression of the DNA sequences of this invention may
25 be isolated from fermentation or animal cell cultures and purified using any of a variety of conventional methods. One of skill in the art may select the most appropriate isolation and purification techniques without departing from the scope of this invention.

30 One can also produce gelsolin fusion polypeptides by chemical synthesis using conventional peptide synthesis techniques, such as solid phase synthesis [R.B. Merrifield, "Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide", J. Am. Chem. Soc., 83, pp. 2149-54 (1963)].
35

- 27 -

Another method useful for producing gelsolin fusion polypeptides, in addition to genetic fusion and chemical synthesis, is by chemically coupling the functional moiety to the gelsolin moiety. This method
5 is useful for both chemical moieties and polypeptide moieties.

Several approaches are available for chemically coupling the gelsolin moiety to a polypeptide moiety. The preferable strategy is to
10 identify or create sites on the polypeptide moiety through which it may be selectively linked to the gelsolin moiety without destroying the activity of the polypeptide moiety. Glycoproteins, such as CD4, have limited numbers of sugars that are useful as cross-
15 linking sites. The sugars may be oxidized to aldehydes and an aldehyde then reacted with an amine group on the gelsolin moiety to create an aldehyde-amine linkage.
[P.K. Nakane and A. Kawaoi, "Peroxidase Labelled Antibody: A New Method of Conjugation", J. Histochem. Cytochem., 22, p. 1084 (1984) and T.-H. Liao et al.,
20 "Modification of Sialyl Residues of Sialolycoprotein(s) of the Human Erythrocyte Surface", J. Biol. Chem., 248, pp. 8247-53 (1973)]. CD4 has two functional glycosylation sites at amino acids +269 to +271 and
25 +298 to +300 (see SEQ ID NO:3). These are outside the gp120 binding region, which is within the first 113 amino acids of the protein [B.H. Chao et al., "A 113-amino Acid Fragment of CD4 Produced in Escherichia coli Blocks Human Immunodeficiency Virus-induced Cell
30 Fusion", J. Biol. Chem., 264, pp. 5812-17 (1989)]. Therefore, coupling CD4 through the carbohydrate should not interfere with function. Alternatively, CD4 may be genetically engineered to eliminate one of the glycosylation sites. This would increase selectivity
35 during linkage. We describe aldehyde-amine linkages in Example II using CD4.

- 28 -

Protein chemists have also developed specific chemistries for covalently coupling polypeptides through thiol groups. A polypeptide moiety having a free thiol may be linked to a gelsolin moiety containing a cysteine either by direct formation of a disulfide bond or indirectly through a homo-bifunctional crosslinker. One example of a homo-bifunctional crosslinker is bismaleimido-hexane (BMH) which has thiol-reactive maleimide groups and forms covalent bonds with free thiols. These methods require the construction of a gelsolin moiety with a cysteine at the amino- or carboxy-terminus. Peptide synthesizers (Example II, Section 2) are useful for in these constructions.

If the polypeptide moiety does not have a free thiol group, such a group may be introduced. For example, the polypeptide may be bound to a thiol-containing amine. More particularly, an oxidized sugar on the polypeptide moiety may be reacted with the amine as described above.

Also, a cysteine may be introduced into the amino acid sequence of the polypeptide moiety by site-directed mutagenesis.

Alternatively, the polypeptide moiety and the gelsolin moiety may be crosslinked through hetero-bifunctional crosslinking agents. These are chosen so that one of the functional groups binds to a group on the polypeptide moiety and the other binds to the thiol on the gelsolin moiety. For example, a succinimide group could bind to an amine group on the polypeptide moiety and a thiol-reactive group, such as a maleimide or an activated thiol could bind to a cysteine on the gelsolin moiety.

We describe methods involving thiol linkage in Example III using CD4. The Pierce Co. Immunotechnology Catalogue and Handbook Volume 1

- 29 -

§§ E4-E12, E41-E48 and E31-E40 describes many useful homo- and hetero-bifunctional crosslinkers, thiol-containing amines and molecules with reactive groups.

Other methods useful for coupling both
5 polypeptide and chemical moieties include, for example, those employing glutaraldehyde [M. Reichlin, "Use of Glutaraldehyde as a Coupling Agent for Proteins and Peptides", Methods In Enzymology, 70, pp. 159-65 (1980)], N-ethyl-N'-(3-dimethylaminopropyl)-
10 carbodiimide [T.L. Goodfriend et al., "Antibodies to Bradykinin and Angiotensin: A Use of Carbodiimides in Immunology", Science, 144, pp. 1344-46 (1964)] or a mixture of N-ethyl-N'-(3-dimethylaminopropyl)-carbodiimide and a succinylated carrier [M.H. Klapper
15 and I.M. Klotz, "Acylation with Dicarboxylic Acid Anhydrides", Methods In Enzymology, 25, pp. 531-36 (1972)]. Since chemical coupling is not limited to one site on the gelsolin moiety, it is possible to couple more than one functional moiety to each gelsolin
20 moiety.

Multimeric and hetero-multimeric gelsolin fusion constructs according to this invention may be produced by binding gelsolin fusion polypeptides to phospholipids aggregated into a vesicle. The vesicle
25 must comprise at least one phospholipid that binds to gelsolin, but may contain others as well. The phosphatidylinositols, PIP and PIP₂, are preferable components of the vesicle because they bind to gelsolin. To be effective the vesicles preferably
30 contain at least 3% of PIP or PIP₂. Other lipids that may comprise the vesicle include, but are not limited to, phosphatidylcholine (PC), phosphatidyl ethanolamine (PE), phosphatidylserine (PS). One may also create vesicles containing detergents such as Triton.

35 The production of phospholipid vesicles is well known to the art [D.M. Haverstick and M. Glaser,

- 30 -

"Visualization of Ca^{2+} -induced Phospholipid Domains",
Proc. Natl. Acad. Sci., USA, 64, pp. 4475-79 (1987)].
For example, dried lipids are mixed with water and the
mixture is sonicated, producing vesicles. PIP should
5 be sonicated more thoroughly than PIP_2 in order to
obtain vesicles of similar size and binding. The
gelsolin fusion polypeptide is then added and allowed
to bind to the vesicles. The resulting product is a
multimeric gelsolin fusion construct.

10 The fact that a vesicle may comprise many
different lipids and detergents allows great
flexibility in engineering a fusion construct with
desired characteristics. For example, one may produce
vesicles that bind different numbers of gelsolin fusion
15 polypeptides by varying the lipid composition of the
starting materials to create larger vesicles, or by
increasing the percentage of PIP or PIP_2 in the vesicle.
Also, one may alter the half-life of the functional
moiety. We expect that these vesicles will be subject
20 to eventual degradation by lipases. By altering the
lipid composition of the vesicle, one could vary the
degradation rate of the vesicle.

When phospholipid vesicles containing
cavities are prepared in the presence of a bioactive
25 molecule, such as those illustrated herein, that
molecule will come to be enclosed within the vesicles.
Accordingly, it is possible to produce a multimeric
gelsolin fusion construct that encloses within it a
bioactive agent. These liposomes may fuse with cell
30 membranes, delivering their contents to cells and
adding the gelsolin fusion polypeptide to the cell
membrane.

Hetero-multimeric gelsolin fusion constructs
comprise at least two different functional moieties or
35 two different gelsolin moieties. For example, hetero-
multimeric gelsolin fusion constructs may comprise two

- 31 -

different polypeptide moieties, two different chemical moieties or both a polypeptide moiety and a chemical moiety.

Hetero-multimeric gelsolin fusion constructs
5 are especially useful when the properties of the different moieties complement one another. For example, it is possible to combine receptors that bind to a particular target particle or cell with toxins or anti-retroviral agents in fusion proteins according to
10 this invention to produce targeted toxic or anti-retroviral agents. Polypeptides useful as toxins include, but are not limited to, ricin, abrin, angiogenin, Pseudomonas Exotoxin A, pokeweed antiviral protein, saponin, gelonin and diphtheria toxin, or
15 toxic portions thereof. Useful anti-retroviral agents include suramin, azidothymidine (AZT), dideoxycytidine and glucosidase inhibitors such as castanospermine, deoxynojirimycin and derivatives thereof.

Hetero-multimeric gelsolin fusion constructs
20 according to this invention are also useful as diagnostic agents to identify the presence of a target molecule in a sample or in vivo. Such proteins comprise one functional moiety which is a recognition molecule, such as an immunoglobulin or a fragment
25 thereof (Fab, dAb) that binds to the target molecule [See Ward et al., supra] and a second functional moiety, which is a reporter group, such as a radionuclide, an enzyme (such as horseradish peroxidase) or a fluorescent or chemiluminescent
30 marker. Typically, the reporter group will be bound directly to the reporter group; for example, HRP is bound directly to the immunoglobulin. Many reporter groups may be coupled to a multimeric gelsolin fusion constructs thereby enhancing the signal. These
35 constructs may be used, for example, to replace antibodies as the recognition molecules that contact

- 32 -

the sample in ELISA-type assays, or as in vivo imaging agents.

Hetero-multimeric gelsolin fusion constructs according to this invention may also be used as multi-
5 vaccines. For example, one may produce such constructs using several different antigenic determinants from the same infective agent. Also, one can produce constructs comprising antigenic determinants from several
10 infective agents, such as polio, measles, mumps and others used for childhood vaccination.

The pharmaceutical compositions of this invention typically comprise a pharmaceutically effective amount of a multimeric gelsolin fusion construct and a pharmaceutically acceptable carrier.
15 Therapeutic methods of this invention comprise the step of treating patients in a pharmaceutically acceptable manner with those compositions. These compositions may be used to treat any mammal, including humans.

The pharmaceutical compositions of this invention may be in a variety of forms. These include,
20 for example, solid, semi-solid and liquid dosage forms, such as tablets, pills, powders, liquid solutions or suspensions, liposomes, suppositories, injectable and infusible solutions and sustained release forms. The
25 preferred form depends on the intended mode of administration and therapeutic application. The compositions also preferably include conventional pharmaceutically acceptable carriers and adjuvants which are known to those of skill in the art.

30 Generally, the pharmaceutical compositions of the present invention may be formulated and administered using methods and compositions similar to those used for pharmaceutically important polypeptides such as, for example, alpha interferon. The fusion
35 constructs of this invention may be administered by conventional routes of administration, such as

- 33 -

parenteral, subcutaneous, intravenous, intramuscular or intralesional routes. It will be understood that conventional doses will vary depending upon the particular molecular moiety involved.

5 In order that this invention may be better understood, the following examples are set forth. These examples are for the purposes of illustration only, and are not to be construed as limiting the scope of the invention in any manner.

10 In the examples that follow, the molecular biology techniques employed, such as cloning, cutting with restriction enzymes, isolating DNA fragments, filling out with Klenow enzyme and deoxyribonucleotides triphosphate (dXTP), ligating, transforming E.coli and
15 the like are conventional protocols exemplified and further described in J. Sambrook et al., Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (1989).

20 EXAMPLE I - PRODUCTION OF A CD4-GELSOLIN FUSION POLYPEPTIDE BY GENETIC FUSION

1. Cloning of pCD4-Gelsolin

We constructed a plasmid expression vector containing a DNA sequence encoding a CD4-gelsolin fusion polypeptide and used it to transform E.coli.
25 The coding region contains a DNA sequence for CD4(181) fused to the 5' end of 140 bp fragment encoding a 12 amino-acid spacer and amino acids 150-173 of gelsolin. This includes the PIP_2 binding domain. We constructed the plasmid as follows. (See Figure 6.)

30 First, we produced a DNA sequence containing the human gelsolin PIP_2 binding domain. The PIP_2 binding domain is encompassed within amino acids +150 to +169 (nucleotides 541-600) of Figure 1 (SEQ ID NO:1). We created this DNA sequence from the plasmid
35 pM1D which contains the cDNA human gelsolin-encoding

- 34 -

sequence of Figure 1 (SEQ ID NO:1). (Plasmid pM1D was the gift of David Kwiatkowski, Harvard Medical School, Boston, Massachusetts.) We amplified a cDNA sequence for the PIP₂ binding domain using polymerase chain
5 reaction (PCR) (Sambrook et al., Chapter 14). We carried out all amplifications using Taq DNA polymerase and primers prephosphorylated with T4 polynucleotide kinase and ATP. We used the oligonucleotide ACE 144 (SEQ ID NO:3) as the sense primer (which hybridizes to
10 the anti-sense strand) and ACE 145 (SEQ ID NO:4) as the anti-sense primer. (See Figure 5.) We filled out the amplified fragments with Klenow enzyme and dXTP. This produced blunt-ended 140 bp DNA fragments having a BglIII site near the 5' end and an EcoRI site near the
15 3' end. The fragments encoded gelsolin amino acids +143 through +173 (see SEQ ID NO:1).

Then we digested an intermediate plasmid, pNN03, with EcoRV and dephosphorylated the ends to prevent recircularization. Plasmid pNN03 is derived
20 from pUC13 by the incorporation of a polylinker. (Pharmacia PL Biochemicals). We subcloned the 140 bp fragments into this plasmid. We called the resulting plasmid pGell.

We then inserted the BalII/EcoRI DNA fragment
25 encoding the gelsolin PIP₂ binding domain from pGell into a prokaryotic expression vector containing a DNA sequence encoding CD4(181) and derived from pEX56.

Plasmid pEX56 encodes CD4(181) fused in-frame to the 5' end of a DNA insert encoding
30 Pseudomonas endotoxin. The insert is bordered by EcoRI sites at the 5' and 3' ends and contains a BglIII site at the junction of the CD4-endotoxin sequence. The Pseudomonas endotoxin gene has been altered to remove the ribosome binding region. Plasmid pEX56 is created
35 by site-directed mutagenesis of pEX46 (Example III, section 2 and Figure 13 (SEQ ID NO:13) with

- 35 -

oligonucleotide T4-AID 176 (Figure 5, SEQ ID NO:9).

[The plasmid is described in co-pending PCT application PCT/US89/04584, incorporated herein by reference.]

We digested a first sample of pEX56 with
5 EcoRI and BglII and isolated the 613 bp fragment that
encodes CD4(181). Then we digested a second sample of
pEX56 with EcoRI, dephosphorylated the fragments, and
isolated the 3922 bp fragment representing the pEX56
vector portion. We ligated together the 3922 bp EcoRI
10 fragment, the 613 bp EcoRI/BglII fragment and the 140
bp BglII/EcoRI fragment. We used this ligation mixture
to transform E.coli JA221 [ATCC 33875] by standard
CaCl₂ procedures. (See Sambrook, Chapter 1.82.) We
identified the plasmids pCD4-gelsolin and p α CD4-
15 gelsolin (opposite orientation and therefore non-
expressing) by restriction digests of mini-plasmid DNA
preparations. The plasmid map of pCD4-gelsolin is
shown in Figure 8. The DNA sequence and predicted
amino acid sequence of the CD4-gelsolin fusion
20 polypeptide obtained is shown in Figure 7 (SEQ ID
NO:10). We have deposited an isolate of pCD4-gelsolin
with In Vitro International, IVI-10253.

2. Expression of CD4-Gelsolin

We transformed E.coli JA221 and E.coli A89
25 (an htpR⁻ protease deficient mutant) with pCD4-gelsolin
and p α CD4-gelsolin. E.coli A89 is a tetracycline-
sensitive mutant of E.coli SG936 [ATCC 39624]. We then
tested the cultures for the production of CD4-gelsolin.
Our results showed that pCD4-gelsolin, but not p α CD4-
30 gelsolin, produced a polypeptide of the molecular
weight predicted for CD4-gelsolin.

We grew 5 ml overnight cultures in LB + 12.5
 μ g/ml tetracycline at 30°C. We diluted the overnight
cultures 1:10 into LB + 12.5 μ g/ml tetracycline and
35 grew the cultures until the optical density at 550 nm

- 36 -

was between 1 and 1.5. We then added the culture to an equal volume of LB + 12.5 μ g/ml tetracycline at 42°C. After two hours we harvested the cells, lysed them, and analyzed the contents for a protein band corresponding to the size expected for a CD4-gelsolin fusion molecule by SDS-polyacrylamide gel electrophoresis (SDS-PAGE). We thus identified a protein having molecular weight of about 28 kD.

We have isolated this protein using the protocol of Example III, section 2b.

EXAMPLE II CHEMICAL CROSS-LINKING OF A GELSOLIN MOIETY TO CD4 VIA ALDEHYDE-AMIDE LINKAGE

We cross-linked CD4(375) (a gift of Biogen, Inc., Cambridge, Massachusetts) to a gelsolin moiety by oxidizing sugars on the CD4 glycoprotein to aldehydes and then reacting an aldehyde with an amine on the gelsolin moiety to create an aldehyde-amine linkage.

1. Oxidation of CD4(375)

We dialyzed 100 μ M CD4(375) protein against 0.1 M sodium acetate pH 5.0 at 4°C. We incubated the preparation at 23°C for 1 hour with 1 mM aqueous sodium periodate and immediately desalted on a P6DG column (BioRad, Richmond, California) that was equilibrated in 10 mM sodium acetate pH 5.0, 100 mM NaCl. We stored the oxidized CD4(375) at 4°C for subsequent use or at -70°C for long term storage. We monitored the extent of oxidation by measuring incorporation of tritiated sodium borohydride. Typically 8-10 aldehydes per CD4(375) were generated.

To confirm that oxidation did not interfere with the CD4(375) function, we assessed the ability of the modified protein to bind gp120 in an ELISA format. We coated IMMULON II® plates (Dynatech Laboratories, Chantilly, Virginia), with gp120 (a gift of Biogen,

- 37 -

Inc., and commercially available from American Bio-Technologies, Inc., Cambridge, Massachusetts), added CD4(375) or oxidized CD4(375), and then determined the binding with a reporter system using
5 OKT4 antibody (available from Ortho Diagnostics Systems, Raritan, New Jersey) that was conjugated with horseradish peroxidase. There was no difference in binding of soluble CD4 protein or oxidized CD4 to gp120. Upon amino acid analysis, both samples were
10 also found to be similar with no apparent effect of oxidation on individual amino acids.

2. Reaction of Oxidized CD4(375)
with the Gelsolin Moiety, GEL1

We synthesized a gelsolin moiety, GEL1, using
15 an Applied Biosystems 430A peptide synthesizer. GEL1 has the amino acid sequence Gly-Tyr-Gly-Lys-His-Val-Val-Pro-Asn-Glu-Val-Val-Val-Gln-Arg-Leu-Phe-Gln-Val-Lys-Gly-Arg-Arg (SEQ ID NO:14). The final twenty amino acids constitute the PIP₂-binding sequence of gelsolin,
20 amino acids +150 to +169 (see SEQ ID NO:1). To crosslink GEL1 with CD4(375), we incubated varying concentrations of GEL1 overnight at 23°C with 10 µM oxidized CD4(375) in the presence of 50 mM MES, pH 6.5, and 5 mM sodium cyanoborohydride.

25 We tested the sample for crosslinking by SDS-PAGE. Samples were either analyzed directly by staining with Coomassie brilliant blue or by Western blotting using an antiserum raised in rabbits against GEL1. The immunogen consisted of GEL1 crosslinked to
30 Keyhole limpet hemocyanin with glutaraldehyde.

We found a dose dependent increase in the molecular weight of CD4 treated with GEL1, indicating that the protein had become modified. At low peptide concentrations, there was little effect on the mobility
35 of CD4(375) but when incubated with 1 mM GEL1,

- 38 -

approximately 50% of the CD4(375) migrated with an increased apparent molecular weight that is consistent with it containing one GEL1 peptide per CD4(375). When CD4(375) was incubated with 10 mM GEL1, all of the protein shifted to a high-molecular weight form. We observed a series of bands that likely correspond to moieties with one, two, and three gelsolin moieties per CD4(375). The need for a large molar excess of GEL1 over CD4(375) to drive the crosslinking reaction is consistent with the results obtained for modifying periodate oxidized CD4 with other amino-containing reagents as well. (See Example III.)

To verify that GEL1 had been crosslinked to CD4(375), we analyzed selected fractions by Western blotting using antibodies against GEL1. A prominent immunoreactive band was observed in the sample after crosslinking. This band is absent from the Western blot of an untreated CD4 sample.

3. Analysis of the CD4-Gelsolin Fusion Polypeptide

We demonstrated above that the crosslinking chemistry did not affect the ability of CD4(375) to bind gp120. We have further established that CD4(375)-gelsolin fusion polypeptides bind to PIP₂ vesicles.

We assayed the ability of CD4(375)-gelsolin to associate with PIP or PIP₂ vesicles using an aggregation assay similar to that described by Janmey et al., "Phosphoinositide Micelles and Polyphosphoinositide-containing Vesicles Dissociate Endogenous Gelsolin-actin Complexes and Promote Actin Assembly From the Fast-growing End of Actin Filaments Blocked by Gelsolin", J. Biol. Chem., 262, pp. 12228-36 (1987). In the assay, the amount of protein used is appropriately adjusted to take into account the molecular weight of the CD4-gelsolin fusion

- 39 -

polypeptide. Mg^{++} causes micelles of pure polyphosphoinositides to aggregate into larger vesicles, increasing the turbidity of the solution. However, gelsolin inhibits this aggregation. We found
5 that CD4(375)-gelsolin behaved like the GEL1 peptide in this assay. Recombinant SCD4, alone, had no activity in this assay.

Because the junction between the gelsolin peptide fragment and the spacer is unnatural, it may be
10 necessary to change the composition or length of the spacer region in order to optimize function. This involves resynthesizing the gelsolin peptide fragment with other sequences added at either the amino or carboxy terminus of the polypeptide. The coupling
15 chemistry would not be affected. Alternatively, it may be advantageous to change selected amino acids from the binding sequence in order to change the affinity of the fusion polypeptide for PIP_2 .

20 EXAMPLE III STRATEGIES FOR CROSSLINKING CD4
THROUGH THIOL GROUPS

We describe here three strategies for crosslinking the CD4 polypeptide moiety with a gelsolin moiety through thiol groups. They involve the modification of the CD4 protein to contain a cysteine,
25 a free thiol or a thiol-reactive group.

1. Introducing a Free Thiol into CD4

First, a thiol group may be introduced into CD4 using thiol-containing amines, such as cysteine, cystamine or glutathione. An aldehyde is introduced
30 into CD4 and then one creates an aldehyde-amine linkage (see Example II). Once the thiol-containing CD4 is generated, it can be selectively crosslinked to the gelsolin moiety.

- 40 -

We incubated periodate oxidized CD4(375) (0.5 mg/ml) overnight at 23°C in 50 mM MES, pH 6.5, 5 mM sodium cyanoborohydride with 20 mM of either cysteine, oxidized cystamine or oxidized glutathione to create CD4(cysteine), CD4(cystamine), and CD4(glutathione). We treated the samples with 40 mM DTT for 40 minutes at 23°C. We then dialyzed them against storage buffer (10 mM sodium acetate, pH 5.0, 100 mM NaCl). We monitored the extent of modification with Ellman's reagent. Briefly, we diluted the samples into 100 μ l of 100 mM sodium phosphate pH 8.0, 0.5 mM DTNB and measured the absorbance after 5 minutes at 410 nm. We calibrated the samples against a standard curve that was developed with reduced glutathione. Both cystamine and glutathione treatments resulted in three to five groups per CD4. For subsequent studies, the preparations were concentrated to 5 mg/ml using a CENTRICON-10® filtration unit (Amicon, Danvers, Massachusetts).

These molecules may be bound to gelsolin moieties through the thiol groups using homo-bifunctional crosslinking agents with two thiol-reactive groups, such as BMH or o- or p-phenylene dimaleimide. We believe that this method will result in crosslinking because treatment of CD4(cystamine) with these agents induced the formation of CD4 dimers and higher molecular weight complexes. With sub-stoichiometric amounts of crosslinker we were able to drive crosslinking of CD4 to greater than 50%. A similar strategy will be used with the cysteine-containing gelsolin moiety where a dimaleimide agent will be used to generate crosslinking complexes.

Alternatively, the moieties may be crosslinked through disulfide bonds using conventional techniques.

- 41 -

2. Introducing a Free Cysteine into
CD4 by Site-Specific Mutagenesis

Second, a free cysteine may be introduced in the primary sequence of CD4 through genetic engineering. Crosslinking to the gelsolin moiety is then directed using the methods of section 1 of this example. We describe herein the construction and isolation of two truncated forms of CD4 engineered to contain cysteine residues at their C-termini.

10 a. Construction of pDC219
and Expression of CD4(111Cys)

To produce CD4(111Cys) we constructed the expression plasmid pDC219. (See Figure 9.) We began with p218-8, a plasmid in which the λP_L promoter controls the expression of CD4(111). This plasmid is described in PCT patent application WO 89/0194, p. 77/93, Figure 28. The DNA sequence for p218-8 is depicted in Figure 10 (SEQ ID NO:11). We digested a first sample of p218-8 with PstI and BglII and isolated the 3645 bp fragment. We then digested a second sample of p218-8 with PstI and EcoRI and isolated the 269 bp fragment. We digested a third sample of p218-8 with EcoRI and BspMI and isolated the 395 bp fragment. We isolated these fragments by electrophoresing the digests on agarose gels, cutting out the relevant bands and electroeluting the DNA fragments. We precipitated the electroeluted DNA fragments with ethanol, centrifuged the mixture to pellet the DNA fragments and resuspended the fragments in 10 mM Tris-HCl, pH 8.0, 1 mM Na₂EDTA.

We phosphorylated oligonucleotides T4AID-133 (SEQ ID NO:5) and T4AID-134 (SEQ ID NO:6) (Figure 5) using bacteriophage T4 polynucleotide kinase. These oligonucleotides contain a BglII recognition sequence.

- 42 -

Then we ligated the purified DNA fragments and the oligonucleotides.

We used the reaction mixture to transform E.coli DH1. We selected colonies that grew at 30°C,
5 12.5 µg/ml tetracycline and analyzed them for the correct sequences by digestion with BglII. We subjected those plasmid DNAs which had the additional BglII site to DNA sequence analysis. Thus we obtained pDC219.

10 To produce CD4(111Cys), we transformed A89 cells with pDC219 and fermented the cells at a 10 liter scale. (We achieved an expression level of 13%.) We stored cells as frozen cell pellets.

To isolate CD4(111Cys) we thawed 50 g frozen
15 whole cells, suspended them in 20 mM Tris pH 7.5, 1 mM EDTA, 0.4 mg/ml lysozyme, and mixed with a Polytron (Brinkman Instruments, Westbury, N.Y.). We stirred the cell slurry at room temperature for one hour, then passed it three times through a prechilled Manton
20 Gaulin French press (550 setting). We chilled the lysate on ice between each passage. We pelleted particulates in a SA600 rotor for 15 minutes at 10,000 rpm. We washed the resulting pellet twice with a 1:4 dilution in 20 mM Tris pH 9.0 and pelleted it as
25 before. (All ratios given are whole cell weight to buffer volume.) We then washed the pellet with a 1:4 dilution in 20 mM Tris pH 9.0 containing 0.5 M NaCl and spun down the pellet using previous conditions by resuspending with a Polytron. We extracted the final
30 pellet in a 1:4 dilution of extraction buffer (7 M urea, 20 mM Tris 9.0, 10 mM β-mercaptoethanol) and stirring at room temperature for 15 minutes. We removed debris by centrifugation in a SA600 rotor at 15,000 G for 30 minutes.

35 We diluted the clarified supernatant 1:4 with fresh extraction buffer and passed it over a Fast S

- 43 -

cation exchange column (Pharmacia) pre-equilibrated with extraction buffer at a column ratio of 1 gm whole cells to 4 ml resin. We washed the column extensively with extraction buffer. We then eluted the protein
5 with salt steps of half column volume of extraction buffer containing 0.05 M, 0.075 M, 0.1 M, 0.15 M and 0.2 M NaCl, respectively. CD4(111Cys) routinely eluted in the 0.15 M NaCl step.

We pooled the CD4(111Cys) peak and diluted it
10 to an absorbance of under O.D. 0.5 at 280 nm. Then we dialyzed the sample overnight, 1:100 V:V, with one change, against 3 M urea, 20 mM Tris pH 7.5. We diluted the dialysate to 1 M urea with the 20 mM Tris pH 7.5, and filtered it through 0.45 μ sterile filter
15 unit. We bound CD4 from the filtrate to 6C6-Sepharose for one hour at 4°C with rocking. 6C6 is a monoclonal antibody developed at Biogen that recognizes CD4 and blocks CD4 binding to gp120. Alternatively, one may use anti-Leu-3a, a monoclonal available from Becton-
20 Dickinson, Mountain View, California. Then we poured the slurry into a column and washed with 2 x 0.5 column volumes 50 mM Tris pH 7.5, 0.5 M NaCl (wash 2), and 2 x 0.5 column volumes of wash 1 buffer (wash 3). CD4(111Cys) was eluted from the resin with a series of
25 0.1 column volume additions of 50 mM glycine, pH 3.0, 250 mM NaCl. We neutralized the eluate by the addition of 2 M Tris pH 9.0 to 50 mM.

The resulting affinity purified protein was 90% CD4(111Cys) monomer with contaminating multimeric
30 bands. When run under reducing conditions these additional bands collapsed into the monomer, indicating they were disulfide forms of the protein. From 1 gm wet weight of cells we recovered between 0.5 to 0.75 mg of CD4(111Cys). We assayed the gp120 binding activity
35 and found it to be about half the specific activity that is observed for full length CD4.

- 44 -

We carried out biotinylation studies using maleimidobutyryl biocytin (MBB) to test the susceptibility of the engineered cysteine to modification with the maleimide. We monitored biotin labeling on Western blots using avidin-conjugated HRP to track the biotin. Specific biotin labeling of CD4(111Cys) was observed when fresh samples were analyzed; however, the efficiency of labeling decreased with time as the samples aged.

10 b. Construction of λ P_L180cys and
 Expression of CD4(180Cys)

To produce CD4(180Cys), we constructed the expression plasmid λ P_L180Cys, in which a λ P_L promoter controls the expression of a DNA sequence encoding
15 CD4(180Cys). (See Figure 11.)

We began with plasmid pBG391, an animal cell expression vector that expresses CD4(375). (The DNA sequence of this plasmid is set forth in Figure 12 (SEQ ID NO:12)). We cleaved pBG391 with StuI. StuI cuts
20 the CD4 gene at the codon for amino acid 182. We phosphorylated oligonucleotides T4AID-137 (SEQ ID NO:7) and T4AID-138 (SEQ ID NO:8) (Figure 5) and ligated into the StuI-cleaved pBG391. This generated pBG398C2. We identified pBG398C2 by the presence of a BamHI site,
25 generated at the junction of the StuI site and T4AID-137.

Then we cleaved pBG398C2 with SacI and BglII and isolated the 490 bp fragment. We cleaved pEX46 with SacI and BamHI and isolated the large fragment.
30 (The DNA sequence of pEX46 is set forth in Figure 13 (SEQ ID NO:13)). Then we ligated the two fragments together. This generated plasmid λ P_L180cys.

In 10 liter fermentations, CD4(180Cys) was expressed at about 5% of the total cell protein.

- 45 -

We suspended fermentation cells at 8 ml/gm cell wet weight in 20 mM Tris-HCl, 1 mM Na₂EDTA, pH 7.7, broke them in two passes through a French press and washed them twice with 20 ml/gm cell wet weight of
5 1 M guanidine-HCl, 1 M urea, 15 mM sodium acetate, pH 5 followed by two washes in 20 mM Tris-HCl, 1 mM Na₂EDTA, pH 7.7. We extracted the washed pellet with 25 ml/gm cell wet weight of 6 M guanidine-HCl, 20 mM Tris-HCl, 10 mM DTT, pH 7.7 overnight at room
10 temperature. We spun the suspension for 45 minutes in a SS-34 rotor at 20,000 rpm. We diluted the supernatant 1:60 into cold 20 mM Tris-HCl, pH 7.7 and added BSA to a final concentration of 0.5 mg/ml.

To generate microgram amounts of the protein,
15 we concentrated the diluted extract by ultrafiltration using a PM10® membrane (Amicon) followed by affinity purification on 6C6-Sepharose 4B. Alternatively, CD4(180Cys) may be prepared as follows: The pH of the diluted extract obtained as described above is lowered
20 to 7.0 with HCl and loaded at 1% vol/vol onto a Fast S column equilibrated in 20 mM Tris-HCl, pH 7.0. Bound protein is washed with 5 column volumes of equilibration buffer and eluted with 0.2 M NaCl in the same buffer. The elution pool is diluted with one
25 volume of 20 mM Tris-HCl, pH 7.7 and loaded on a 6C6-Sepharose 4B column. The bound protein is washed and eluted from the affinity column in 50 mM glycine, 250 mM NaCl, pH 3.0. The elution fractions are neutralized with 1/15 volume of 0.5 M HEPES pH 7.5,
30 pooled according to the A₂₈₀ profile and stored at 4°C.

One may bind CD4(111cys) or CD4(180cys) to a thiol-containing gelsolin moiety using the chemistries described in section 1 of this example.

- 46 -

3. Hetero-bifunctional Crosslinking Agents

According to a third method, CD4 may also be crosslinked with a cysteine-containing gelsolin moiety using a hetero-bifunctional crosslinking agent. Such crosslinkers include succinimidyl 4-(N-maleimidomethyl) cyclohexane-1-carboxylate (SMCC), m-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS), or N-succinimidyl 3-(2-pyridyldithiol) proprionate (SPDP). The succinimidyl arms of these crosslinkers bind to primary amines in CD4. The reactive thiol (maleimide) of SMCC and MBS and the activated thiol of SPDP react with the thiol from the cysteine in the gelsolin moiety to form the covalent linkage.

To carry out the reaction with SMCC and MBS, the crosslinker is incubated with CD4 for 0.5 hours at pH 6.0 at 23°C. Unreacted crosslinker is then removed on a desalting column. SPDP is used as described in the Pharmacia Co. Users Manual. A gelsolin moiety having a free terminal cysteine is then added. The mixture is incubated for 3 hours at 23°C, creating the covalent linkage. Unreacted gelsolin moiety is removed on a desalting column.

The extent and specificity of the modification can be analyzed as described in Example II. The lysine content of CD4 is high; therefore reactions with lysine would not provide much specificity. However, by limiting the amount of crosslinker added, it may be possible to direct crosslinking to one or a small number of lysines that are particularly reactive.

Alternatively, one may bind the reactive thiol group of the hetero-bifunctional crosslinker to a thiol group introduced into CD4 and then bind the succinimidyl arm to an amine in the gelsolin moiety.

- 47 -

EXAMPLE IV - MULTIMERIC GELSOLIN FUSION CONSTRUCT

We have shown that CD4-gelsolin fusion polypeptides retain affinity for gp120 and that they bind PIP₂ vesicles through the gelsolin moiety. This demonstrates that the chemistry we have developed to produce multimeric gelsolin fusion constructs is sound. As a next step, we produced and tested a multimeric CD4(375)-gelsolin fusion construct.

Multimeric gelsolin fusion constructs comprising CD4-gelsolin fusion polypeptides were produced using methods that involve binding the fusion polypeptides to PIP₂ vesicles.

PIP₂ vesicles were produced in the following manner. PIP₂ may be obtained as a lyophilized solid (Sigma Chemical Co., St. Louis, Missouri). Water was added to the dried sample to a concentration of 1 to 3 mg/ml and the mixture was sonicated for between 30 seconds to 2 minutes at maximum intensity in a Heat Systems - Ultrasonics, Inc. (Farmingdale, New York) W185® apparatus or its equivalent until an optically clear solution formed. These samples were kept at 4°C and used within a week or they were stored frozen for future use. For storage, the samples were divided into aliquots, frozen in liquid nitrogen and stored at -70° until use. Prior to use, the samples were thawed quickly under a stream of warm water and sonicated for 30 minutes at room temperature in a water bath sonicator.

CD4-gelsolin fusion polypeptides were then added to lipid at a 5 to 10 molar excess of lipid over protein and the mixture was incubated at room temperature for about five minutes.

We tested the ability of the multimeric CD4(375)-gelsolin fusion construct to bind gp120 in an ELISA-type assay. Briefly, we coated plates with

- 48 -

gp120, added the fusion construct and assayed for binding using anti-CD4 as the reporter antibody. We did not detect binding of the multimeric CD4(375)-gelsolin fusion construct to gp120.

5 We also tested the biological activity of the fusion construct in a viral replication assay similar to the one described in co-pending United States application 07/583,022 (incorporated herein by reference). Briefly, we incubated the fusion construct
10 with HIV, added cells from a T-cell line, and measured the incidence of infection. Multimeric CD4(375)-gelsolin fusion construct did not block infection in this assay.

 As a result, we found that rsCD4, itself,
15 binds to PIP₂ vesicles and that in doing so, its ability to bind gp120 is inactivated. Recombinant sCD4 has pockets of positive charge that cause it to bind to cation exchange matrices with high avidity at neutral pH. Since PIP₂ vesicles, like cation exchange
20 matrices, possess high negative charge, we believe that the binding of rsCD4 to PIP₂ vesicles is due to its ionic character.

 Therefore, one may produce multimeric CD4-gelsolin fusion constructs that bind gp120 by altering
25 the charge of the CD4 moiety so that it no longer binds PIP₂ vesicles. The first one-hundred-thirteen amino acids of rsCD4, which contain the gp120 binding domain, contain sixteen basic amino acid residues: thirteen lysine residues and three arginine residues. Using
30 site specific mutagenesis, one may alter one or more of these into histidine, a basic, but less polar amino acid, or into neutral amino acids. Among these alternate versions of CD4, one may select molecules that bind gp120 but do not bind PIP₂ vesicles. We
35 believe that these alternate versions of CD4 would be

- 49 -

useful to produce multimeric CD4-gelsolin fusion constructs that possess gp120 binding ability.

Although they do not bind gp120, multimeric CD4(181)-gelsolin fusion constructs have other uses.

5 For example, they are useful as immunogens to elicit α -CD4 antibodies. In diagnostic assays, they are useful to detect the presence of α -CD4 in a sample. A percentage of patients infected with HIV exhibit α -CD4 antibodies.

10 Positive charge at neutral pH and high salt concentration is uncommon among proteins. Accordingly, we do not believe that many proteins other than CD4 would exhibit deactivation when employed to produce multimeric-gelsolin fusion constructs according to this
15 invention. Nevertheless, the ionic character and lipid-binding properties of potential functional moieties are factors to be considered in predicting the ultimate biological activity and characteristics of multimeric gelsolin fusion constructs produced using
20 them.

Microorganisms and recombinant DNA molecules according to this invention are exemplified by cultures deposited in the In Vitro International, Inc. culture collection, in Linthicum, Maryland, USA on May 4, 1990,
25 and identified as:

pCD4-gelsolin	IVI-10253
p170.2	IVI-10252.

While we have hereinbefore described a number of embodiments of this invention, it is apparent that
30 our basic embodiments can be altered to provide other embodiments which utilize the processes and compositions of this invention. Therefore, it will be appreciated that the scope of this invention includes all alternative embodiments and variations which are
35 defined in the foregoing specification and by the

- 50 -

claims appended hereto; and the invention is not to be limited by the specific embodiments which have been presented herein by way of example.

- 51 -

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: PEPINSKY, R. BLAKE
ROSA, MARGARET D.
STOSSEL, THOMAS P.
- (ii) TITLE OF INVENTION: MULTIMERIC GELSOLIN FUSION CONSTRUCTS
- (iii) NUMBER OF SEQUENCES: 14
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: FISH & NEAVE
 - (B) STREET: 875 Third Avenue
 - (C) CITY: New York
 - (D) STATE: New York
 - (E) COUNTRY: United States of America
 - (F) ZIP: 10022
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER:
 - (B) FILING DATE:
 - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 07/520,368
 - (B) FILING DATE: 04-MAY-1990
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Haley Jr., James F.
 - (B) REGISTRATION NUMBER: 27,794
 - (C) REFERENCE/DOCKET NUMBER: B144CIP
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: (212) 715-0600
 - (B) TELEFAX: (212) 715-0634
 - (C) TELEX: 14-8367

- 52 -

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2588 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ATGGCTCCGC ACCGCCCCGC GCCCGCGCTG CTTTGCGCGC TGTCCCTGGC GCTGTGCGCG	60
CTGTGCTGCTGC CCGTCCGCGC GGCCACTGCG TCGCGGGGGG CGTCCCAGGC GGGGGCGCCC	120
CAGGGGCGGG TGCCCGAGGC GCGGCCCAAC AGCATGGTGG TGGAACACCC CGAGTTCCTC	180
AAGGCAGGGA AGGAGCCTGG CCTGCAGATC TGGCGTGTGG AGAAGTTCGA TCTGGTGCCC	240
GTGCCCACCA ACCTTTATGG AGACTTCTTC ACGGGCGACG CCTACGTCAT CCTGAAGACA	300
GTGCAGCTGA GGAACGGAAA TCTGCAGTAT GACCTCCACT ACTGGCTGGG CAATGAGTGC	360
AGCCAGGATG AGAGCGGGGC GGCCGCCATC TTTACCGTGC AGCTGGATGA CTACCTGAAC	420
GGCCGGGGCG TGCAGCACCG TGAGGTCCAG GGCTTCGAGT CGGCCACCTT CCTAGGCTAC	480
TTCAAGTCTG GCCTGAAGTA CAAGAAAGGA GGTGTGGCAT CAGGATTCAA GCACGTGGTA	540
CCCAACGAGG TGGTGGTGCA GAGACTCTTC CAGGTCAAAG GGCGGCGTGT GGTCCGTGCC	600
ACCGAGGTAC CTGTGTCCTG GGAGAGCTTC AACAAATGGCG ACTGCTTCAT CCTGGACCTG	660
GGCAACAACA TCCACCAGTG GTGTGGTTCC AACAGCAATC GGTATGAAAG ACTGAAGGCC	720
ACACAGGTGT CCAAGGGCAT CCGGGACAAC GAGCGGAGTG GCCGGGCCCC AGTGCACGTG	780
TCTGAGGAGG GCACTGAGCC CGAGGCGATG CTCCAGGTGC TGGGCCCCAA GCCGGCTCTG	840
CCTGCAGGTA CCGAGGACAC CGCCAAGGAG GATGCGGCCA ACCGCAAGCT GGCCAAGCTC	900
TACAAGGTCT CCAATGGTGC AGGGACCATG TCCGTCTCCC TCGTGGCTGA TGAGAACCCC	960
TTCGCCCAGG GGGCCCTGAA GTCAGAGGAC TGCTTCATCC TGGACCACGG CAAAGATGGG	1020
AAAATCTTTG TCTGGAAAGG CAAGCAGGCA AACACGGAGG AGAGGAAGGC TGCCCTCAAA	1080
ACAGCCTCTG ACTTCATCAC CAAGATGGAC TACCCCAAGC AGACTCAGGT CTCGGTCCTT	1140
CCTGAGGGCG GTGAGACCCC ACTGTTCAAG CAGTTCTTCA AGAACTGGCG GGACCCAGAC	1200
CAGACAGATG GCCTGGGCTT GTCCTACCTT TCCAGCCATA TCGCCAACGT GGAGCGGGTG	1260

- 53 -

CCCTTCGACG	CCGCCACCCT	GCACACCTCC	ACTGCCATGG	CCGCCCAGCA	CGGCATGGAT	1320
GACGATGGCA	CAGGCCAGAA	ACAGATCTGG	AGAATCGAAG	GTTCCAACAA	GGTGCCCGTG	1380
GACCCTGCCA	CATATGGACA	GTTCTATGGA	GGCGACAGCT	ACATCATTCT	GTACAACTAC	1440
CGCCATGGTG	GCCGCCAGGG	GCAGATAATC	TATAACTGGC	AGGGTGCCCA	GTCTACCCAG	1500
GATGAGGTG	CTGCATCTGC	CATCCTGACT	GCTCAGCTGG	ATGAGGAGCT	GGGAGGTACC	1560
CCTGTCCAGA	GCCGTGTGGT	CCAAGGCAAG	GAGCCCGCCC	ACCTCATGAG	CCTGTTTGGT	1620
GGGAAGCCCA	TGATCATCTA	CAAGGGCGGC	ACCTCCCGCG	AGGGCGGGCA	GACAGCCCCT	1680
GCCAGCACCC	GCCTCTTCCA	GGTCCGCGCC	AACAGCGCTG	GAGCCACCCG	GGCTGTTGAG	1740
GTATTGCCTA	AGGCTGGTGC	ACTGAACTCC	AACGATGCCT	TTGTTCTGAA	AACCCCTCA	1800
GCCGCCTACC	TGTGGGTGGG	TACAGGAGCC	AGCGAGGCAG	AGAAGACGGG	GGCCCAGGAG	1860
CTGCTCAGGG	TGCTGCGGGC	CCAACCTGTG	CAGGTGGCAG	AAGGCAGCGA	GCCAGATGGC	1920
TTCTGGGAGG	CCCTGGGCGG	GAAGGCTGCC	TACCGCACAT	CCCCACGGCT	GAAGGACAAG	1980
AAGATGGATG	CCCATCCTCC	TCGCCTCTTT	GCCTGCTCCA	ACAAGATTGG	ACGTTTTGTG	2040
ATCGAAGAGG	TTCCTGGTGA	GCTCATGCAG	GAAGACCTGG	CAACGGATGA	CGTCATGCTT	2100
CTGGACACCT	GGGACCAGGT	CTTTGTCTGG	GTTGGAAAGG	ATTCTCAAGA	AGAAGAAAAG	2160
ACAGAAGCCT	TGACTTCTGC	TAAGCGGTAC	ATCGAGACGG	ACCCAGCCAA	TCGGGATCGG	2220
CGGACGCCCA	TCACCGTGGT	GAAGCAAGGC	TTTGAGCCTC	CCTCCTTTGT	GGGCTGGTTC	2280
CTTGGCTGGG	ATGATGATTA	CTGGTCTGTG	GACCCCTTGG	ACAGGGCCAT	GGCTGAGCTG	2340
GCTGCCTGAG	GAGGGGCAGG	GCCCACCCAT	GTCACCGGTC	AGTGCCTTTT	GGAAGTGTCC	2400
TTCCCTCAAA	GAGGCCTTAG	AGCGAGCAGA	GCAGCTCTGC	TATGAGTGTG	TGTGTGTGTG	2460
TGTGTTGTTT	CTTTTTTTTT	TTTTTACAGT	ATCCAAAAAT	AGCCCTGCAA	AAATTCAGAG	2520
TCCTTGCAAA	ATTGTCTAAA	ATGTCAGTGT	TTGGGAAATT	AAATCCAATA	AAAACATTTT	2580
GAAGTGTG						2588

- 54 -

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1377 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

ATGAACCGGG GAGTCCCTTT TAGGCACTTG CTTCTGGTGC TGCAACTGGC GTCCTCCCA	60
GCAGCCACTC AGGGAAAGAA AGTGGTGCTG GGCAAAAAG GGGATACAGT GGAAGTACC	120
TGTACAGCTT CCCAGAAGAA GAGCATACAA TTCCACTGGA AAAACTCCAA CCAGATAAAG	180
ATTCTGGGAA ATCAGGGCTC CTTCTTAACT AAAGGTCCAT CCAAGCTGAA TGATCGCGCT	240
GACTCAAGAA GAAGCTTGTTG GGACCAAGGA AACTTTCCCC TGATCATCAA GAATCTTAAG	300
ATAGAAGACT CAGATACTTA CATCTGTGAA GTGGAGGACC AGAAGGAGGA GGTGCAATTG	360
CTAGTGTTTCG GATTGACTGC CAACTCTGAC ACCCACCTGC TTCAGGGGCA GAGCCTGACC	420
CTGACCTTGG AGAGCCCCC TGGTAGTAGC CCCTCAGTGC AATGTAGGAG TCCAAGGGGT	480
AAAAACATAC AGGGGGGGAA GACCCTCTCC GTGTCTCAGC TGGAGCTCCA GGATAGTGGC	540
ACCTGGACAT GCACTGTCTT GCAGAACCAG AAGAAGGTGG AGTTCAAAT AGACATCGTG	600
GTGCTAGCTT TCCAGAAGGC CTCCAGCATA GTCTACAAGA AAGAGGGGGA ACAGGTGGAG	660
TTCTCCTTCC CACTCGCCTT TACAGTTGAA AAGCTGACGG GCAGTGGCGA GCTGTGGTGG	720
CAGGCGGAGA GGGCTTCCTC CTCCAAGTCT TGGATCACCT CTGACCTGAA GAACAAGGAA	780
GTGTCTGTAA AACGGGTTAC CCAGGACCCT AAGCTCCAGA TGGGCAAGAA GCTCCCGCTC	840
CACCTCACCC TGCCCCAGGC CTTGCCTCAG TATGCTGGCT CTGGAAACCT CACCCTGGCC	900
CTTGAAGCGA AAACAGGAAA GTTGCATCAG GAAGTGAACC TGGTGGTGAT GAGAGCCACT	960
CAGCTCCAGA AAAATTTGAC CTGTGAGGTG TGGGGACCCA CCTCCCCTAA GCTGATGCTG	1020
AGCTTGAAAC TGGAGAACAA GGAGGCAAAG GTCTCGAAGC GGGAGAAGGC GGTGTGGGTG	1080
CTGAACCCTG AGGCGGGGAT GTGGCAGTGT CTGCTGAGTG ACTCGGGACA GGTCTGCTG	1140
GAATCCAACA TCAAGGTTCT GCCCACATGG TCGACCCCGG TGCAGCCAAT GGCCCTGATT	1200
GTGCTGGGGG GCGTCGCCGG CCTCCTGCTT TTCATTGGGC TAGGCATCTT CTTCTGTGTC	1260

- 55 -

AGGTGCCGGC ACCGAAGGCG CCAAGCAGAG CGGATGTCTC AGATCAAGAG ACTCCGCAGT 1320

GAGAAGAAGA CCTGCCAGTG CCCTCACCGG TTTCAGAAGA CATGTAGCCC CATTTGA 1377

- 56 -

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 37 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

AGATCTACGG GGGCGTGGCA TCAGGATTCA AGCACGT

37

- 57 -

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 27 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GAATTCTTAG GCACGGACCA CACGCCG

27

- 58 -

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 25 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GGGGTGTGTA TAGTAAGATC TTGCA

25

- 59 -

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 17 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

AGATCTTACT ATCAAGA

17

- 60 -

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 26 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

ATCCCTGTCC GTAGAAGCTT ATCGAT

26

- 61 -

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 26 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

ATCGATAAGC TTCTACGGAC AGGGAT

26

- 62 -

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 46 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GGAGGACCAG AAAGAAGAAG TTCAGCTGCT GGTTTTCGGA TTGACT

46

- 63 -

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 654 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

ATGAAAAAAG TAGTACTGGG CAAAAAAGGG GATACAGTGG AACTGACCTG TACAGCTTCC	60
CAGAAGAAGA GCATACAATT CCACTGGAAA AACTCCAACC AGATAAAGAT TCTGGGAAAT	120
CAGGGCTCCT TCTTAACTAA AGGTCCATCC AAGCTGAATG ATCGCGCTGA CTCAAGAAGA	180
AGCTTGTTGG ACCAAGGAAA CTTTCCCCTG ATCATCAAGA ATCTTAAGAT AGAAGACTCA	240
GATACTTACA TCTGTGAAGT GGAGGACCAG AAAGAAGAAG TTCAGCTGCT GGTTTTCGGA	300
TTGACTGCCA ACTCTGACAC CCACCTGCTT CAGGGGCAGA GCCTGACCCT GACCTTGGAG	360
AGCCCCCTG GTAGTAGCCC CTCAGTGCAA TGTAGGAGTC CAAGGGGTAA AAACATACAG	420
GGGGGGAAGA CCCTCTCCGT GTCTCAGCTG GAGCTCCAGG ATAGTGGCAC CTGGACATGC	480
ACTGTCTTGC AGAACCAGAA GAAGGTGGAG TTCAAAATAG ACATCGTGGT GCTAGCTTTC	540
CAGAAGGGGA AGATCTACGG GGGCGTGGCA TCAGGATTCA AGCACGTGGT ACCCAACGAG	600
GTGGTGGTGC AGAGACTCTT CCAGGTCAAA GGGCGGCGTG TGGTCCGTGC CTAA	654

- 64 -

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4309 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GAATTCTTAC ACTTAGTTAA ATTGCTAACT TTATAGATTA CAAACTTAG GAAATCGATT	60
TGGATGAAAA AAGTAGTACT GGGCAAAAAA GGGGATACAG TGGAAGTGAC CTGTACAGCT	120
TCCCAGAAGA AGAGCATACA ATTCCACTGG AAAAACTCCA ACCAGATAAA GATTCTGGGA	180
AATCAGGGCT CCTTCTTAAC TAAAGGTCCA TCCAAGCTGA ATGATCGCGC TGA CTCAAGA	240
AGAAGCTTGT GGGACCAAGG AACTTTCCC CTGATCATCA AGAATCTTAA GATAGAAGAC	300
TCAGATACTT ACATCTGTGA AGTGGAGGAC CAGAAGGAGG AGGTGCAATT GCTAGTGTTT	360
GGATTGACTG CCAACTCTGA CACCCACCTG CTCAGGGGT GATAGTAAGA TCCTGCAGCC	420
CAGCTTGGGG ACCCTAGAGG TCCCCTTTTT TATTTTGAAT TGGGAGATCC CAATTCTCAT	480
GTTTGACAGC TTATCATCGA TAAGCTAGCT TTAATGCGGT AGTTTATCAC AGTTAAATTG	540
CTAACGCAGT CAGGCACCGT GTATGAAATC TAACAATGCG CTCATCGTCA TCCTCGGCAC	600
CGTCACCCTG GATGCTGTAG GCATAGGCTT GGTATGCCG GTAGTGCCG GCCTCTTGCG	660
GGATATCGTC CATTCCGACA GCATCGCCAG TCACTATGGC GTGCTGCTAG CGCTATATGC	720
GTTGATGCAA TTTCTATGCG CACCCGTTCT CGGAGCACTG TCCGACCGCT TTGGCCGCCG	780
CCCAGTCCTG CTCGCTTCGC TACTTGAGC CACTATCGAC TACGCGATCA TGGCGACCAC	840
ACCCGTCCTG TGGATTCTCT ACGCCGGACG CATCGTGGCC GGCATCACCG GCGCCACAGG	900
TGCGGTTGCT GCGCCCTATA TCGCCGACAT CACCGATGGG GAAGATCGGG CTCGCCACTT	960
CGGGCTCATG AGCGCTTGTT TCGGCGTGGG TATGGTGGCA GGGCCCGTGG CCGGGGGACT	1020
GTTGGGCGCC ATCTCCTTGC ACGCACCATT CCTTGCGGCG GCGGTGCTCA ACGGCCTCAA	1080
CCTACTACTG GGCTGCTTCC TAATGCAGGA GTCGCATAAG GGAGAGCGTC GTCCGATGCC	1140
CTTGAGAGCC TTCAACCCAG TCAGCTCCTT CCGGTGGGCG CGGGGCATGA CTATCGTCGC	1200
CGCACTTATG ACTGTCTTCT TTATCATGCA ACTCGTAGGA CAGGTGCCGG CAGCGCTCTG	1260

- 65 -

GGTCATTTTC	GGCGAGGACC	GCTTTCGCTG	GAGCGCGACG	ATGATCGGCC	TGTCGCTTGC	1320
GGTATTCGGA	ATCTTGACAG	CCCTCGCTCA	AGCCTTCGTC	ACTGGTCCCG	CCACCAAACG	1380
TTTCGGCGAG	AAGCAGGCCA	TTATCGCCGG	CATGGCGGCC	GACGCGCTGG	GCTACGTCTT	1440
GCTGGCGTTC	GCGACGCGAG	GCTGGATGGC	CTTCCCCATT	ATGATTCTTC	TCGCTTCCGG	1500
CGGCATCGGG	ATGCCCCGCT	TGCAGGCCAT	GCTGTCCAGG	CAGGTAGATG	ACGACCATCA	1560
GGGACAGCTT	CAAGGATCGC	TCGCGGCTCT	TACCAGCCTA	ACTTCGATCA	CTGGACCGCT	1620
GATCGTCACG	GCGATTTATG	CCGCCTCGGC	GAGCACATGG	AACGGGTTGG	CATGGATTGT	1680
AGGCGCCGCC	CTATACCTTG	TCTGCCTCCC	CGCGTTGCGT	CGCGGTGCAT	GGAGCCGGGC	1740
CACCTCGACC	TGAATGGAAG	CCGGCGGCAC	CTCGCTAACG	GATTCACCAC	TCCAAGAATT	1800
GGAGCCAATC	AATTCTTGCG	GAGAACTGTG	AATGCGCAAA	CCAACCCTTG	GCAGAACATA	1860
TCCATCGCGT	CCGCCATCTC	CAGCAGCCGC	ACGCGGCGCA	TCTCGGGGGA	TGATCAGCTG	1920
CCTCGCGCGT	TTCGGTGATG	ACGGTGAAAA	CCTCTGACAC	ATGCAGCTCC	CGGAGACGGT	1980
CACAGCTTGT	CTGTAAGCGG	ATGCCGGGAG	CAGACAAGCC	CGTCAGGGCG	CGTCAGCGGG	2040
TGTTGGCGGG	TGTCGGGGCG	CAGCCATGAC	CCAGTCACGT	AGCGATAGCG	GAGTGTATAC	2100
TGGCTTAACT	ATGCGGCATC	AGAGCAGATT	GTACTGAGAG	TGCACCATAT	GCGGTGTGAA	2160
ATACCGCACA	GATGCGTAAG	GAGAAAATAC	CGCATCAGGC	GCTCTTCCGC	TTCCTCGCTC	2220
ACTGACTCGC	TGCGCTCGGT	CGTTCGGCTG	CGGCGAGCGG	TATCAGCTCA	CTCAAAGGCG	2280
GTAATACGGT	TATCCACAGA	ATCAGGGGAT	AACGCAGGAA	AGAACATGTG	AGCAAAAGGC	2340
CAGCAAAAGG	CCAGGAACCG	TAAAAAGGCC	GCGTTGCTGG	CGTTTTTCCA	TAGGCTCCGC	2400
CCCCCTGACG	AGCATCACAA	AAATCGACGC	TCAAGTCAGA	GGTGGCGAAA	CCCGACAGGA	2460
CTATAAAGAT	ACCAGGCGTT	TCCCCCTGGA	AGCTCCCTCG	TGCGCTCTCC	TGTTCCGACC	2520
CTGCCGCTTA	CCGATACCT	GTCCGCCTTT	CTCCCTTCGG	GAAGCGTGGC	GCTTTCTCAA	2580
TGCTCACGCT	GTAGGTATCT	CAGTTCGGTG	TAGGTCGTTT	GCTCCAAGCT	GGGCTGTGTG	2640
CACGAACCCC	CCGTTACAGC	CGACCGCTGC	GCCTTATCCG	GTAACATATC	TCTTGAGTCC	2700
AACCCGGTAA	GACACGACTT	ATCGCCACTG	GCAGCAGCCA	CTGGTAACAG	GATTAGCAGA	2760
GCGAGGTATG	TAGGCGGTGC	TACAGAGTTC	TTGAAGTGGT	GGCCTAACTA	CGGCTACACT	2820
AGAAGGACAG	TATTTGGTAT	CTGCGCTCTG	CTGAAGCCAG	TTACCTTCGG	AAAAAGAGTT	2880

- 66 -

GGTAGCTCTT	GATCCGGCAA	ACAAACCACC	GCTGGTAGCG	GTGGTTTTTT	TGTTTGCAAG	2940
CAGCAGATTA	CGCGCAGAAA	AAAAGGATCT	CAAGAAGATC	CTTTGATCTT	TTCTACGGGG	3000
TCTGACGCTC	AGTGGAACGA	AAACTCACGT	TAAGGGATTT	TGGTCATGAG	ATTATCAAAA	3060
AGGATCTTCA	CCTAGATCCT	TTTCAGATCT	CCCGATCTTT	AGCTGTCTTG	GTTTGCCCAA	3120
AGCGCATTGC	ATAATCTTTC	AGGGTTATGC	GTTGTTCCAT	ACAACCTCCT	TAGTACATGC	3180
AACCATTATC	ACCGCCAGAG	GTAAAATAGT	CAACACGCAC	GGTGTTAGAT	ATTTATCCCT	3240
TGCGGTGATA	GATTTAACGT	ATGAGCACAA	AAAAGAAACC	ATTAACACAA	GAGCAGCTTG	3300
AGGACGCACG	TCGCCTTAAA	GCAATTTATG	AAAAAAGAA	AAATGAACTT	GGCTTATCCC	3360
AGGAATCTGT	CGCAGACAAG	ATGGGGATGG	GGCAGTCAGG	CGTTGGTGCT	TTATTTAATG	3420
GCATCAATGC	ATTAAATGCT	TATAACGCCG	CATTGCTTAC	AAAAATTCTC	AAAGTTAGCG	3480
TTGAAGAATT	TAGCCCTTCA	ATCGCCAGAG	AAATCTACGA	GATGTATGAA	GCGGTTAGTA	3540
TGCAGCCGTC	ACTTAGAAGT	GAGTATGAGT	ACCCTGTTTT	TTCTCATGTT	CAGGCAGGGA	3600
TGTTCTCACC	TAAGCTTAGA	ACCTTTACCA	AAGGTGATGC	GGAGAGATGG	GTAAGCACAA	3660
CCAAAAAAGC	CAGTGATTCT	GCATTCTGGC	TTGAGGTTGA	AGGTAATTCC	ATGACCGCAC	3720
CAACAGGCTC	CAAGCCAAGC	TTTCCTGACG	GAATGTTAAT	TCTCGTTGAC	CCTGAGCAGG	3780
CTGTTGAGCC	AGGTGATTTT	TGCATAGCCA	GACTTGGGGG	TGATGAGTTT	ACCTTCAAGA	3840
AACTAATTAG	GGATAGCGGT	CAGGTGTTTT	TACAACCACT	AAACCCACAG	TACCCAATGA	3900
TCCCATGCAA	TGAGAGTTGT	TCCGTTGTGG	GGAAAGTTAT	CGCTAGTCAG	TGGCCTGAAG	3960
AGACGTTTGG	CTGATCGGCA	AGGTGTTCTG	GTCGGCGCAT	AGCTGATAAC	AATTGAGCAA	4020
GAATCTTCAT	CGGGGCTGCA	GCCCACGATG	CGTCCGGCGT	AGAGGATCTC	TCACCTACCA	4080
AACAATGCCC	CCCTGCAAAA	AATAAATTCA	TATAAAAAAC	ATACAGATAA	CCATCTGCGG	4140
TGATAAATTA	TCTCTGGCGG	TGTTGACATA	AATACCACTG	GCGGTGATAC	TGAGCACATC	4200
AGCAGGACGC	ACTGACCACC	ATGAAGGTGA	CGCTCTTAAA	ATTAAGCCCT	GAAGAAGGGC	4260
AGCATTCAAA	GCAGAAGGCT	TTGGGGTGTG	TGATACGAAA	CGAAGCATT		4309

- 67 -

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6151 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GAATTAATTC CAGCTTGCTG TGAATGTGT GTCAGTTAGG GTGTGGAAAG TCCCCAGGCT	60
CCCCAGCAGG CAGAAGTATG CAAAGCATGC ATCTCAATTA GTCAGCAACC AGGTGTGGAA	120
AGTCCCCAGG CTCCCCAGCA GGCAGAAGTA TGCAAAGCAT GCATCTCAAT TAGTCAGCAA	180
CCATAGTCCC GCCCCTAACT CCGCCCATCC CGCCCCTAAC TCCGCCCAGT TCCGCCCATT	240
CTCCGCCCCA TGGCTGACTA ATTTTTTTTA TTTATGCAGA GGCCGAGGCC GCCTCGGCCT	300
CTGAGCTATT CCAGAAGTAG TGAGGAGGCT TTTTGGAGG GGTCTCCTC GTATAGAAAC	360
TCGGACCACT CTGAGACGAA GGCTCGCGTC CAGGCCAGCA CGAAGGAGGC TAAGTGGGAG	420
GGGTAGCGGT CGTTGTCCAC TAGGGGGTCC ACTCGCTCCA GGGTGTGAAG ACACATGTCG	480
CCCTCTTCGG CATCAAGGAA GGTGATTGGT TTATAGGTGT AGGCCACGTG ACCGGGTGTT	540
CCTGAAGGGG GGCTATAAAA GGGGGTGGGG GCGCGTTCGT CCTCACTCTC TTCCGCATCG	600
CTGTCTGCGA GGGCCAGCTG TTGGGCTCGC GGTTGAGGAC AACTCTTCG CGGTCTTTCC	660
AGTACTCTTG GATCGGAAAC CCGTCGGCCT CCGAACGGTA CTCCGCCACC GAGGGACCTG	720
AGCGAGTCCG CATCGACCGG ATCGGAAAAC CTCTCGAGAA AGGCGTCTAA CCAGTCACAG	780
TCGCAAGGTA GGCTGAGCAC CGTGGCGGGC GGCAGCGGGT GGCGGTCGGG GTTGTTTCTG	840
GCGGAGGTGC TGCTGATGAT GTAATTAAAG TAGGCGGTCT TGAGACGGCG GATGGTCGAG	900
GTGAGGTGTG GCAGGCTTGA GATCGATCTG GCCATACACT TGAGTGACAA TGACATCCAC	960
TTTGCCTTTC TCTCCACAGG TGTCCACTCC CAGGTCCAAC TGGATCCAAG CTTCGACTCG	1020
AGGAATTCCC CGAAGGAACA AAGCACCCCTC CCCACTGGGC TCCTGGTTGC AGAGCTCCAA	1080
GTCCTCACAC AGATACGCCCT GTTTGAGAAG CAGCGGGCAA GAAAGACGCA AGCCCAGAGG	1140
CCCTGCCATT TCTGTGGGCT CAGGTCCCTA CTGGCTCAGG CCCCTGCCTC CCTCGGCAAG	1200
GCCACAATGA ACCGGGGAGT CCCTTTTAGG CACTTGCTTC TGGTGCTGCA ACTGGCGCTC	1260

- 68 -

CTCCCAGCAG	CCACTCAGGG	AAAGAAAGTG	GTGCTGGGCA	AAAAAGGGGA	TACAGTGGAA	1320
CTGACCTGTA	CAGCTTCCCA	GAAGAAGAGC	ATACAATTCC	ACTGGAAAAA	CTCCAACCAG	1380
ATAAAGATTG	TGGGAAATCA	GGGCTCCTTC	TTAACTAAAG	GTCCATCCAA	GCTGAATGAT	1440
CGCGCTGACT	CAAGAAGAAG	CTTGTGGGAC	CAAGGAAACT	TTCCCCTGAT	CATCAAGAAT	1500
CTTAAGATAG	AAGACTCAGA	TACTTACATC	TGTGAAGTGG	AGGACCAGAA	GGAGGAGGTG	1560
CAATTGCTAG	TGTTCGGATT	GACTGCCAAC	TCTGACACCC	ACCTGCTTCA	GGGGCAGAGC	1620
CTGACCCTGA	CCTTGGAGAG	CCCCCCTGGT	AGTAGCCCCCT	CAGTGCAATG	TAGGAGTCCA	1680
AGGGGTAAAA	ACATACAGGG	GGGGAAGACC	CTCTCCGTGT	CTCAGCTGGA	GCTCCAGGAT	1740
AGTGGCACCT	GGACATGCAC	TGTCTTGACG	AACCAGAAGA	AGGTGGAGTT	CAAAATAGAC	1800
ATCGTGGTGC	TAGCTTTCCA	GAAGGCCTCC	AGCATAGTCT	ATAAGAAAGA	GGGGGAACAG	1860
GTGGAGTTCT	CCTTCCCCT	CGCCTTTACA	GTTGAAAAGC	TGACGGGCAG	TGGCGAGCTG	1920
TGGTGGCAGG	CGGAGAGGGC	TTCCTCCTCC	AAGTCTTGGA	TCACCTTTGA	CCTGAAGAAC	1980
AAGGAAGTGT	CTGTAAAACG	GGTTACCCAG	GACCCTAAGC	TCCAGATGGG	CAAGAAGCTC	2040
CCGCTCCACC	TCACCCTGCC	CCAGGCCTTG	CCTCAGTATG	CTGGCTCTGG	AAACCTCACC	2100
CTGGCCCTTG	AAGCGAAAAC	AGGAAAGTTG	CATCAGGAAG	TGAACCTGGT	GGTGATGAGA	2160
GCCACTCAGC	TCCAGAAAAA	TTTGACCTGT	GAGGTGTGGG	GACCCACCTC	CCCTAAGCTG	2220
ATGCTGAGTT	TGAAACTGGA	GAACAAGGAG	GCAAAGGTCT	CGAAGCGGGA	GAAGGCGGTG	2280
TGGGTGCTGA	ACCCTGAGGC	GGGGATGTGG	CAGTGTCTGC	TGAGTGACTC	GGGACAGGTC	2340
CTGCTGGAAT	CCAACATCAA	GGTTCTGCCC	ACATGGTCGA	CCCCGGTGCA	GCCAATGGCC	2400
CTGATTTGAG	ATCTTTGTGA	AGGAACCTTA	CTTCTGTGGT	GTGACATAAT	TGGACAAACT	2460
ACCTACAGAG	ATTTAAAGCT	CTAAGGTAAA	TATAAAATTT	TTAAGTGTAT	AATGTGTTAA	2520
ACTACTGATT	CTAATTGTTT	GTGTATTTTA	GATTCCAACC	TATGGAACTG	ATGAATGGGA	2580
GCAGTGGTGG	AATGCCTTTA	ATGAGGAAAA	CCTGTTTTGC	TCAGAAGAAA	TGCCATCTAG	2640
TGATGATGAG	GCTACTGCTG	ACTCTCAACA	TTCTACTCCT	CCAAAAAGA	AGAGAAAGGT	2700
AGAAGACCCC	AAGGACTTTC	CTTCAGAATT	GCTAAGTTTT	TTGAGTCATG	CTGTGTTTAG	2760
TAATAGAACT	CTTGCTTGCT	TTGCTATTTA	CACCACAAAG	GAAAAAGCTG	CACTGCTATA	2820
CAAGAAAATT	ATGGAAAAAT	ATTCTGTAAC	CTTTATAAGT	AGGCATAACA	GTTATAATCA	2880

- 69 -

TAACATACTG	TTTTTTCTTA	CTCCACACAG	GCATAGAGTG	TCTGCTATTA	ATAACTATGC	2940
TCAAAAATTG	TGTACCTTTA	GCTTTTTTAAT	TTGTAAAGGG	GTTAATAAGG	AATATTTGAT	3000
GTATAGTGCC	TTGACTAGAG	ATCATAATCA	GCCATACCAC	ATTTGTAGAG	GTTTTACTTG	3060
CTTTAAAAAA	CCTCCCACAC	CTCCCCCTGA	ACCTGAAACA	TAAAATGAAT	GCAATTGTTG	3120
TTGTAACTT	GTTTATTGCA	GCTTATAATG	GTTACAAATA	AAGCAATAGC	ATCACAAATT	3180
TCACAAATAA	AGCATTTTTT	TCCTGTCATT	CTAGTTGTGG	TTTGTCCAAA	CTCATCAATG	3240
TATCTTATCA	TGTCTGGATC	CTCTACGCCG	GACGCATCGT	GGCCGGCATC	ACCGGCGCCA	3300
CAGGTGCGGT	TGCTGGCGCC	TATATCGCCG	ACATCACCGA	TGGGGAAGAT	CGGGCTCGCC	3360
ACTTCGGGCT	CATGAGCGCT	TGTTTCGGCG	TGGGTATGGT	GGCAGGCCCG	TGGCCGGGGG	3420
ACTGTTGGGC	GCCATCTCCT	TGCATGCACC	ATTCCTTGCG	GCGGCGGTGC	TCAACGGCCT	3480
CAACCTACTA	CTGGGCTGCT	TCCTAATGCA	GGAGTCGCAT	AAGGGAGAGC	GTCGACCGAT	3540
GCCCTTGAGA	GCCTTCAACC	CAGTCAGCTC	CTTCCGGTGG	GCGCGGGGCA	TGACTATCGT	3600
CGCCGCACTT	ATGACTGTCT	TCTTTATCAT	GCAACTCGTA	GGACAGGTGC	CGGCAGCGCT	3660
CTGGGTCATT	TTCGGCGAGG	ACCGCTTTTCG	CTGGAGCGCG	ACGATGATCG	GCCTGTCGCT	3720
TGCGGTATTC	GGAATCTTGC	ACGCCCTCGC	TCAAGCCTTC	GTCACTGGTC	CCGCCACCAA	3780
ACGTTTCGGC	GAGAAGCAGG	CCATTATCGC	CGGCATGGCG	GCCGACGCGC	TGGGCTACGT	3840
CTTGCTGGCG	TTCGCGACGC	GAGGCTGGAT	GGCCTTCCCC	ATTATGATTC	TTCTCGCTTC	3900
CGGCGGCATC	GGGATGCCCC	CGTTGCAGGC	CATGCTGTCC	AGGCAGGTAG	ATGACGACCA	3960
TCAGGGACAG	CTTCAAGGAT	CGCTCGCGGC	TCTTACCAGC	CTAACTTCGA	TCCTGGACC	4020
GCTGATCGTC	ACGGCGATTT	ATGCCGCCCTC	GGCGAGCACA	TGGAACGGGT	TGGCATGGAT	4080
TGTAGGCGCC	GCCCTATACC	TTGTCTGCCT	CCCCGCGTTG	CGTCGCGGTG	CATGGAGCCG	4140
GGCCACCTCG	ACCTGAATGG	AAGCCGGCGG	CACCTCGCTA	ACGGATTCAC	CACTCCAAGA	4200
ATTGGAGCCA	ATCAATTCTT	GCGGAGAACT	GTGAATGCGC	AAACCAACCC	TTGGCAGAAC	4260
ATATCCATCG	CGTCCGCCAT	CTCCAGCAGC	CGCACGCGGC	GCATCTCGGG	CCGCGTTGCT	4320
GGCGTTTTTC	CATAGGCTCC	GCCCCCTGA	CGAGCATCAC	AAAAATCGAC	GCTCAAAGTCA	4380
GAGGTGGCGA	AACCCGACAG	GACTATAAAG	ATACCAGGCG	TTTCCCCCTG	GAAGCTCCCT	4440
CGTGCGCTCT	CCTGTTCCGA	CCCTGCCGCT	TACCGGATAC	CTGTCCGCCT	TTCTCCCTTC	4500

- 70 -

GGGAAGCGTG	GCGCTTTCTC	AATGCTCACG	CTGTAGGTAT	CTCAGTTCGG	TGTAGGTCGT	4560
TCGCTCCAAG	CTGGGCTGTG	TGCACGAACC	CCCCGTTTCTCAG	CCCGACCGCT	GCGCCTTATC	4620
CGGTAACAT	CGTCTTGAGT	CCAACCCGGT	AAGACACGAC	TTATCGCCAC	TGGCAGCAGC	4680
CACTGGTAAC	AGGATTAGCA	GAGCGAGGTA	TGTAGGCGGT	GCTACAGAGT	TCTTGAAGTG	4740
GTGGCCTAAC	TACGGCTACA	CTAGAAGGAC	AGTATTTGGT	ATCTGCGCTC	TGCTGAAGCC	4800
AGTTACCTTC	GGAAAAAGAG	TTGGTAGCTC	TTGATCCGGC	AAACAAACCA	CCGCTGGTAG	4860
CGGTGGTTTT	TTTGTTTGCA	AGCAGCAGAT	TACGCGCAGA	AAAAAAGGAT	CTCAAGAAGA	4920
TCCTTTGATC	TTTTCTACGG	GGTCTGACGC	TCAGTGGAAC	GAAAACTCAC	GTTAAGGGAT	4980
TTTGGTCATG	AGATTATCAA	AAAGGATCTT	CACCTAGATC	CTTTTAAATT	AAAAATGAAG	5040
TTTTAAATCA	ATCTAAAGTA	TATATGAGTA	AACTTGGTCT	GACAGTTACC	AATGCTTAAT	5100
CAGTGAGGCA	CCTATCTCAG	CGATCTGTCT	ATTTCTGTTCA	TCCATAGTTG	CCTGACTCCC	5160
CGTCGTGTAG	ATAACTACGA	TACGGGAGGG	CTTACCATCT	GGCCCCAGTG	CTGCAATGAT	5220
ACCGCGAGAC	CCACGCTCAC	CGGCTCCAGA	TTTATCAGCA	ATAAACCAGC	CAGCCGGAAG	5280
GGCCGAGCGC	AGAAGTGGTC	CTGCAACTTT	ATCCGCCTCC	ATCCAGTCTA	TTAATTGTTG	5340
CCGGGAAGCT	AGAGTAAGTA	GTTCCGCCAGT	TAATAGTTTG	CGCAACGTTG	TTGCCATTGC	5400
TGCAGGCATC	GTGGTGTAC	GCTCGTCGTT	TGGTATGGCT	TCATTACAGT	CCGGTTCCCA	5460
ACGATCAAGG	CGAGTTACAT	GATCCCCCAT	GTTGTGCAAA	AAAGCGGTTA	GCTCCTTCGG	5520
TCCTCCGATC	GTTGTCAGAA	GTAAGTTGGC	CGCAGTGTTA	TCACTCATGG	TTATGGCAGC	5580
ACTGCATAAT	TCTCTTACTG	TCATGCCATC	CGTAAGATGC	TTTTCTGTGA	CTGGTGAGTA	5640
CTCAACCAAG	TCATTCTGAG	AATAGTGTAT	GCGGCGACCG	AGTTGCTCTT	GCCCCGCGTC	5700
AACACGGGAT	AATACCGCGC	CACATAGCAG	AACTTTAAAA	GTGCTCATCA	TTGGAAAACG	5760
TTCTTCGGGG	CGAAAACTCT	CAAGGATCTT	ACCGCTGTTG	AGATCCAGTT	CGATGTAACC	5820
CACTCGTGCA	CCCAACTGAT	CTTCAGCATC	TTTTACTTTT	ACCAGCGTTT	CTGGGTGAGC	5880
AAAAACAGGA	AGGCAAAATG	CCGCAAAAAA	GGGAATAAGG	GCGACACGGA	AATGTTGAAT	5940
ACTCATACTC	TTCTTTTTTC	AATATTATTG	AAGCATTTAT	CAGGGTTATT	GTCTCATGAG	6000
CGGATACATA	TTTGAATGTA	TTTAGAAAAA	TAAACAAATA	GGGGTTCCGC	GCACATTTCC	6060
CCGAAAAGTG	CCACCTGACG	TCTAAGAAAC	CATTATTATC	ATGACATTAA	CCTATAAAAA	6120

WO 91/17170

PCT/US91/02954

- 71 -

TAGGCGTATC ACGAGGCCCT TTCGTCTTCA A

6151

- 72 -

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5727 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

GAATTCTTAC ACTTAGTTAA ATTGCTAACT TTATAGATTA CAAAACTTAG GAAATCGATT	60
TGGATGAAAA AAGTAGTACT GGGCAAAAAA GGGGATACAG TGGAAGTGAC CTGTACAGCT	120
TCCCAGAAGA AGAGCATACA ATTCCACTGG AAAAACTCCA ACCAGATAAA GATTCTGGGA	180
AATCAGGGCT CCTTCTTAAC TAAAGGTCCA TCCAAGCTGA ATGATCGCGC TGAAGCAAGA	240
AGAAGCTTGT GGGACCAAGG AAACCTTCCC CTGATCATCA AGAATCTTAA GATAGAAGAC	300
TCAGATACTT ACATCTGTGA AGTGGAGGAC CAGAAGGAGG AGGTGCAATT GCTAGTGTTC	360
GGATTGACTG CCAACTCTGA CACCCACCTG CTTCAGGGGC AGAGCCTGAC CCTGACCTTG	420
GAGAGCCCCC CTGGTAGTAG CCCCTCAGTG CAATGTAGGA GTCCAAGGGG TAAAAACATA	480
CAGGGGGGGA AGACCTCTC CGTGTCTCAG CTGGAGCTCC AGGATAGTGG CACCTGGACA	540
TGCACTGTCT TGCAGAACCA GAAGAAGGTG GAGTTCAAAA TAGACATCGT GGTGCTAGCT	600
TTCCAGAAGG GGAAGATCTT TCCCAGGGGC GGCAGCCTGG CCGCGCTGAC CGCGCACCAG	660
GCTTGCCACC TGCCGCTGGA GACTTTCACC CGTCATCGCC AGCCGCGCGG CTGGGAACAA	720
CTGGAGCAGT GCGGCTATCC GGTGCAGCGG CTGGTCGCCC TCTACCTGGC GGCAGCGGCTG	780
TCGTGGAACC AGGTCGACCA GGTGATCCGC AACGCCCTGG CCAGCCCCGG CAGCGGCGGC	840
GACCTGGGCG AAGCGATCCG CGAGCAGCCG GAGCAGGCC GTCTGGCCCT GACCCTGGCC	900
GCCGCCGAGA GCGAGCGCTT CGTCCGGCAG GGCACCGGCA ACGACGAGGC CGGCGCGGCC	960
AACGCCGACG TGGTGAGCCT GACCTGCCCC GTGCGCGCCG GTGAATGCGC GGGCCCGGCG	1020
GACAGCGGCG ACGCCCTGCT GGAGCGCAAC TATCCCACTG GCGCGGAGTT CCTCGGCGAC	1080
GGCGGCGACG TCAGCTTCAG CACCCGCGGC ACGCAGAACT GGACGGTGGG GCGGCTGCTC	1140
CAGGCGCACC GCCAACTGGA GGAGCGCGGC TATGTGTTTG TCGGCTACCA CGGCACCTTC	1200
CTCGAAGCGG CGCAAAGCAT CGTCTTCGGC GGGGTGCGCG CGCGCAGCCA GGACCTCGAC	1260

- 73 -

GCGATCTGGC GCGGTTTCTA TATCGCCGGC GATCCGGCGC TGGCCTACGG CTACGCCCAG	1320
GACCAGGAAC CCGACGCACG CGGCCGGATC CGCAACGGTG CCCTGCTGCG GGTCTATGTG	1380
CCGCGCTCGA GCCTGCCGGG CTTCTACCGC ACCAGCCTGA CCCTGGCCGC GCCGGAGGCG	1440
GCGGGCGAGG TCGAACGGCT GATCGGCCAT CCGCTGCCGC TGC GCCTGGA CGCCATCACC	1500
GGCCCCGAGG AGGAAGGCGG GCGCCTGGAG ACCATTCTCG GCTGGCCGCT GGCCGAGCGC	1560
ACCGTGGTGA TTCCCTCGGC GATCCCCACC GACCCGCGCA ACGTCGGCGG CGACCTCGAC	1620
CCGTCCAGCA TCCCCGACAA GGAACAGGCG ATCAGCGCCC TGCCGGACTA CGCCAGCCAG	1680
CCCGGCAAAC CGCGCGCGA GGACCTGAAG TAACTGCCGC GACCGGCCGG CTCCCTTCGC	1740
AGGAGCCGGC CTTCTCGGGG CCTGGCCATA CATCAGGTTT TCCTGATGCC AGCCCAATCG	1800
AATATGAATT CTCATCGATT TCCATGGGAT CCTGCAGCCC AGCTTGGGGA CCCTAGAGGT	1860
CCCCTTTTTT ATTTTTTGAA TTGGGAGATC CAATTCTCAT GTTTGACAGC TTATCATCGA	1920
AGCTAGCTTT AATGCGGTAG TTTATCACAG TTAAATTGCT AACGCAGTCA GGCACCGTGT	1980
ATGAAATCTA ACAATGCGCT CATCGTCATC CTCGGCACCG TCACCCTGGA TGCTGTAGGC	2040
ATAGGCTTGG TTATGCCGGT ACTGCCGGG CTCTTGCGGG ATATCGTCCA TTCCGACAGC	2100
ATCGCCAGTC ACTATGGCGT GCTGCTAGCG CTATATGCGT TGATGCAATT TCTATGCGCA	2160
CCCGTTCTCG GAGCACTGTC CGACCGCTTT GGCCGCCGCC CAGTCCTGCT CGCTTCGCTA	2220
CTTGAGCCA CTATCGACTA CGCGATCATG GCGACCACAC CCGTCCTGTG GATTCTCTAC	2280
GCCGGACGCA TCGTGGCCGG CATCACCGGC GCCACAGGTG CGGTTGCTGG CGCCTATATC	2340
GCCGACATCA CCGATGGGGA AGATCGGGCT CGCCACTTCG GGCTCATGAG CGCTTGTTTC	2400
GGCGTGGGTA TGGTGGCAGG CCCCCTGGCC GGGGGACTGT TGGGCGCCAT CTCCTTGAC	2460
GCACCATTC TCGCGGCGGC GGTGCTCAAC GGCCCTCAACC TACTACTGGG CTGCTTCCTA	2520
ATGCAGGAGT CGCATAAGGG AGAGCGTCGT CCGATGCCCT TGAGAGCCTT CAACCCAGTC	2580
AGTCCTTCC GGTGGGCGCG GGGCATGACT ATCGTCGCCG CACTTATGAC TGTCTTCTTT	2640
ATCATGCAAC TCGTAGGACA GGTGCCGGCA GCGCTCTGGG TCATTTTCGG CGAGGACCGC	2700
TTTCGCTGGA GCGCGACGAT GATCGGCCTG TCGCTTGCGG TATTCGGAAT CTTGCACGCC	2760
CTCGCTCAAG CCTTCGTCAC TGGTCCCGCC ACCAAACGTT TCGGCGAGAA GCAGGCCATT	2820
ATCGCCGGCA TGGCGGCCGA CGCGCTGGGC TACGTCTTGC TGGCGTTCGC GACGCGAGGC	2880

- 74 -

TGGATGGCCT	TCCCCATTAT	GATTCTTCTC	GCTTCCGGCG	GCATCGGGAT	GCCCCGCTTG	2940
CAGGCCATGC	TGTCCAGGCA	GGTAGATGAC	GACCATCAGG	GACAGCTTCA	AGGATCGCTC	3000
GCGGCTCTTA	CCAGCCTAAC	TTCGATCACT	GGACCGCTGA	TCGTCACGGC	GATTTATGCC	3060
GCCTCGGCGA	GCACATGGAA	CGGGTTGGCA	TGGATTGTAG	GCGCCGCCCT	ATACCTTGTC	3120
TGCCTCCCCG	CGTTGCGTCG	CGGTGCATGG	AGCCGGGCCA	CCTCGACCTG	AATGGAAGCC	3180
GGCGGCACCT	CGCTAACGGA	TTCACCACTC	CAAGAATTGG	AGCCAATCAA	TTCTTGCGGA	3240
GAACTGTGAA	TGCGCAAACC	AACCCTTGGC	AGAACATATC	CATCGCGTCC	GCCATCTCCA	3300
GCAGCCGCAC	GCGGCGCATC	TCGGGGGATG	ATCAGCTGCC	TCGCGCGTTT	CGGTGATGAC	3360
GGTGAAAACC	TCTGACACAT	GCAGCTCCCG	GAGACGGTCA	CAGCTTGTCT	GTAAGCGGAT	3420
GCCGGGAGCA	GACAAGCCCG	TCAGGGCGCG	TCAGCGGGTG	TTGGCGGGTG	TCGGGGCGCA	3480
GCCATGACCC	AGTCACGTAG	CGATAGCGGA	GTGTATACTG	GCTTAACTAT	GCGGCATCAG	3540
AGCAGATTGT	ACTGAGAGTG	CACCATATGC	GGTGTGAAAT	ACCGCACAGA	TGCGTAAGGA	3600
GAAAATACCG	CATCAGGCGC	TCTTCCGCTT	CCTCGCTCAC	TGACTCGCTG	CGCTCGGTCTG	3660
TTCGGCTGCG	GCGAGCGGTA	TCAGCTCACT	CAAAGGCGGT	AATACGGTTA	TCCACAGAAT	3720
CAGGGGATAA	CGCAGGAAAG	AACATGTGAG	CAAAAGGCCA	GCAAAGGCC	AGGAACCGTA	3780
AAAAGGCCGC	GTGCTGGCG	TTTTTCCATA	GGCTCCGCCC	CCCTGACGAG	CATCACAAAA	3840
ATCGACGCTC	AAGTCAGAGG	TGGCGAAACC	CGACAGGACT	ATAAAGATAC	CAGGCGTTTC	3900
CCCCTGGAAG	CTCCCTCGTG	CGCTCTCCTG	TTCCGACCCT	GCCGCTTACC	GGATACCTGT	3960
CCGCCTTTCT	CCCTTCGGGA	AGCGTGGCGC	TTTCTCAATG	CTCACGCTGT	AGGTATCTCA	4020
GTTGCGTGTA	GGTCGTTTCG	TCCAAGCTGG	GCTGTGTGCA	CGAACCCCCC	GTTGAGCCCG	4080
ACCGCTGCGC	CTTATCCGGT	AACTATCGTC	TTGAGTCCAA	CCCGGTAAGA	CACGACTTAT	4140
CGCCACTGGC	AGCAGCCACT	GGTAACAGGA	TTAGCAGAGC	GAGGTATGTA	GGCGGTGCTA	4200
CAGAGTTCTT	GAAGTGGTGG	CCTAACTACG	GCTACACTAG	AAGGACAGTA	TTTGGTATCT	4260
GCGCTCTGCT	GAAGCCAGTT	ACCTTCGGAA	AAAGAGTTGG	TAGCTCTTGA	TCCGGCAAAC	4320
AAACCACCGC	TGGTAGCGGT	GGTTTTTTTTG	TTTGCAAGCA	GCAGATTACG	CGCAGAAAAA	4380
AAGGATCTCA	AGAAGATCCT	TTGATCTTTT	CTACGGGGTC	TGACGCTCAG	TGGAACGAAA	4440
ACTCACGTTA	AGGGATTTTG	GTCATGAGAT	TATCAAAAAG	GATCTTCACC	TAGATCCTTT	4500

- 75 -

TCAGATCTCC CGATCTTTAG CTGTCTTGGT TTGCCCAAAG CGCATTGCAT AATCTTTCAG	4560
GGTTATGCGT TGTTCATAC AACCTCCTTA GTACATGCAA CCATTATCAC CGCCAGAGGT	4620
AAAATAGTCA ACACGCACGG TGTTAGATAT TTATCCCTTG CGGTGATAGA TTTAACGTAT	4680
GAGCACAAAA AAGAAACCAT TAACACAAGA GCAGCTTGAG GACGCACGTC GCCTTAAAGC	4740
AATTTATGAA AAAAAGAAAA ATGAACTTGG CTTATCCCAG GAATCTGTCG CAGACAAGAT	4800
GGGGATGGGG CAGTCAGGCG TTGGTGCTTT ATTTAATGGC ATCAATGCAT TAAATGCTTA	4860
TAACGCCGCA TTGCTTACAA AAATTCTCAA AGTTAGCGTT GAAGAATTTA GCCCTTCAAT	4920
CGCCAGAGAA ATCTACGAGA TGTATGAAGC GGTTAGTATG CAGCCGTCAC TTAGAAGTGA	4980
GTATGAGTAC CCTGTTTTTT CTCATGTTCA GGCAGGGATG TTCTCACCTA AGCTTAGAAC	5040
CTTTACCAAA GGTGATGCGG AGAGATGGGT AAGCACAACC AAAAAAGCCA GTGATTCTGC	5100
ATTCTGGCTT GAGGTTGAAG GTAATTCCAT GACCGCACCA ACAGGCTCCA AGCCAAGCTT	5160
TCCTGACGGA ATGTTAATTC TCGTTGACCC TGAGCAGGCT GTTGAGCCAG GTGATTTCTG	5220
CATAGCCAGA CTTGGGGGTG ATGAGTTTAC CTTCAAGAAA CTAATTAGGG ATAGCGGTCA	5280
GGTGTTTTTA CAACCACTAA ACCCACAGTA CCCAATGATC CCATGCAATG AGAGTTGTTC	5340
CGTTGTGGGG AAAGTTATCG CTAGTCAGTG GCCTGAAGAG ACGTTTGGCT GATCGGCAAG	5400
GTGTTCTGGT CGGCGCATAG CTGATAACAA TTGAGCAAGA ATCTTCATCG GGGCTGCAGC	5460
CCACGATGCG TCCGGCGTAG AGGATCTCTC ACCTACCAAA CAATGCCCCC CTGCAAAAAA	5520
TAAATTCATA TAAAAACAT ACAGATAACC ATCTGCGGTG ATAAATTATC TCTGGCGGTG	5580
TTGACATAAA TACCACTGGC GGTGATACTG AGCACATCAG CAGGACGCAC TGACCACCAT	5640
GAAGGTGACG CTCTTAAAT TAAGCCCTGA AGAAGGGCAG CATTCAAAGC AGAAGGCTTT	5700
GGGGTGTGTG ATACGAAACG AAGCATT	5727

- 76 -

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Gly	Tyr	Gly	Lys	His	Val	Val	Pro	Asn	Glu	Val	Val	Val	Gln	Arg	Leu
1				5					10					15	

Phe	Gln	Val	Lys	Gly	Arg	Arg
				20		

- 77 -

CLAIMS

We claim:

1. A recombinant DNA molecule comprising a DNA sequence encoding a gelsolin fusion polypeptide comprising a first DNA sequence encoding a polypeptide moiety and a second DNA sequence comprising a gelsolin moiety.
2. The recombinant DNA molecule according to claim 1, wherein the gelsolin moiety is derived from human plasma gelsolin.
3. The recombinant DNA molecule according to claim 2, wherein the gelsolin moiety comprises amino acids +1 to +169 of Figure 1 (SEQ ID NO:1).
4. The recombinant DNA molecule according to claim 3, wherein the gelsolin moiety comprises amino acids +150 to +169 of Figure 1 (SEQ ID NO:1).
5. The recombinant DNA molecule according to claim 1, wherein the polypeptide moiety is selected from the group consisting of viral receptors, cell receptors, cell ligands, bacterial immunogens, parasitic immunogens, viral immunogens, immunoglobulins or fragments thereof that bind to target molecules, enzymes, enzyme inhibitors, enzyme substrates, cytokines, growth factors, colony stimulating factors, hormones and toxins.
6. The recombinant DNA molecule according to claim 5, wherein the polypeptide moiety is a soluble CD4 protein.

- 78 -

7. The recombinant DNA molecule according to claim 6, wherein the soluble CD4 protein is selected from the group consisting of CD4(111), CD4(111Cys), CD4(180cys), CD4(181), CD4(183), CD4(187), CD4(345) and CD4(375).

8. The recombinant DNA molecule according to claim 7 which is pCD4-gelsolin.

9. The recombinant DNA molecule according to claim 5, wherein the polypeptide moiety is a cell receptor or a cell ligand selected from the group consisting of ICAM1, ELAM1, VCAM1, VCAM1b, LFA3, CDX and VLA4.

10. The recombinant DNA molecule according to claim 1, wherein the DNA sequence encoding a gelsolin fusion polypeptide is operatively linked to an expression control sequence.

11. The recombinant DNA molecule according to claim 10, wherein the expression control sequence is selected from the group consisting of the early and late promoters of SV40 or adenovirus, the lac system, the trp system, the TAC or TRC system, the major operator and promoter regions of phage λ , the control regions of fd coat protein, the promoter for 3-phosphoglycerate kinase or other glycolytic enzymes, the promoters of acid phosphatase, the promoters of the yeast α -mating factors, the polyhedron promoter of the baculovirus system and other sequences known to control the expression of genes of prokaryotic or eukaryotic cells or their viruses, and various combinations thereof.

- 79 -

12. A recombinant DNA molecule comprising a DNA sequence encoding a lipid binding protein fusion polypeptide comprising a first DNA sequence encoding a polypeptide moiety and a second DNA sequence encoding a lipid binding protein moiety.

13. The recombinant DNA molecule according to claim 12, wherein the lipid binding protein moiety is selected from the group consisting of protein kinase C, lipocortin, severin, villin, fragmin, profilin, cofilin, Cap42(a), gCap39, Cap2, destrin and DNase I.

14. A unicellular host transformed with a recombinant DNA molecule according to claim 1 or 12.

15. The unicellular host according to claim 14, selected from the group consisting of E.coli, Pseudomonas, Bacillus, Streptomyces, fungi, such as yeasts, and animal cells, such as CHO and mouse cells, African green monkey cells, such as COS-1, COS-7, BSC 1, BSC 40, and BMT 10, insect cells, and human cells and plant cells in tissue culture.

16. The unicellular host according to claim 15, said host being a COS-7 cell or a CHO cell.

17. A lipid binding protein fusion polypeptide comprising a functional moiety and a lipid binding protein moiety.

18. The lipid binding protein fusion protein according to claim 17, wherein the lipid binding protein is selected from the group consisting of villin, severin, fragmin, profilin, cofilin, Cap42(a), gCap39, Cap2 and destrin.

- 80 -

19. The lipid binding protein fusion polypeptide according to claim 17, wherein the lipid binding protein is selected from the group consisting of protein kinase C, lipocortin and DNase I.

20. A gelsolin fusion polypeptide comprising a functional moiety and a gelsolin moiety.

21. The gelsolin fusion polypeptide according to claim 20, wherein the functional moiety is a polypeptide moiety.

22. The gelsolin fusion polypeptide according to claim 20, wherein said functional moiety is selected from the group consisting of viral receptors, cell receptors, cell ligands, bacterial immunogens, parasitic immunogens, viral immunogens, immunoglobulins or fragments of them that bind to target molecules, enzymes, enzyme inhibitors, enzyme substrates, cytokines, growth factors, colony stimulating factors, hormones and toxins.

23. The gelsolin fusion polypeptide according to claim 22, wherein said functional moiety is a soluble CD4 protein.

24. The gelsolin fusion polypeptide according to claim 23, wherein the soluble CD4 protein is selected from the group consisting of CD4(111), CD4(111cys) CD4(180cys) CD4(181), CD4(183), CD4(187), CD4(345), CD4(375), CD4(Cystamine), CD4(Cysteine) and CD4(Glutathione).

25. The gelsolin fusion polypeptide according to claim 22, wherein said functional moiety is a cell receptor or a cell ligand selected from the

- 81 -

group consisting of ICAM1, ELAM1, VCAM1, VCAM1b, LFA3, CDX and VLA4.

26. The gelsolin fusion polypeptide according to claim 21, wherein the C-terminus of the polypeptide moiety is fused to the N-terminus of the gelsolin moiety.

27. The gelsolin fusion polypeptide according to claim 21, wherein the polypeptide moiety is chemically coupled to the gelsolin moiety.

28. The gelsolin fusion polypeptide according to claim 27, wherein the polypeptide moiety is chemically coupled to the gelsolin moiety through an aldehyde-amine linkage.

29. The gelsolin fusion polypeptide according to claim 27, wherein the polypeptide moiety is chemically coupled to the gelsolin moiety through a thiol group.

30. The gelsolin fusion polypeptide according to claim 27, wherein the polypeptide moiety comprises an amino-terminal or carboxy-terminal cysteine.

31. The gelsolin fusion polypeptide according to claim 20, wherein said functional moiety is selected from the group consisting of toxins, anti-retroviral agents, enzyme substrates and enzyme inhibitors.

32. The gelsolin fusion polypeptide according to claim 31, wherein the functional moiety is AZT.

- 82 -

33. The gelsolin fusion polypeptide according to claim 20, comprising a reporter group selected from the group consisting of enzymes, radionuclides, fluorescent markers and chemiluminescent markers.

34. A gelsolin fusion construct comprising a gelsolin fusion polypeptide and a vesicle comprising a polyphosphoinositide, said construct being multimeric or hetero-multimeric.

35. The gelsolin fusion construct according to claim 34, wherein the polyphosphoinositide is PIP or PIP₂.

36. The gelsolin fusion construct according to claim 35, said construct comprising a CD4-gelsolin fusion polypeptide.

37. The gelsolin fusion construct according to claim 36, wherein said CD4-gelsolin fusion polypeptide is CD4(181)-gelsolin fusion polypeptide.

38. The gelsolin fusion construct according to claim 34, selected from the group consisting of ELAM1-gelsolin fusion polypeptides, VCAM1-gelsolin fusion polypeptides, VCAM1b-gelsolin fusion polypeptides, ICAM1-gelsolin fusion polypeptides, CDX-gelsolin fusion polypeptides, VLA4-gelsolin fusion polypeptides and LFA3-gelsolin fusion polypeptides.

39. The hetero-multimeric gelsolin fusion construct according to claim 34, said construct comprising a first functional moiety selected from the group consisting of viral receptors, cell receptors and

- 83 -

cell ligands, and a second functional moiety selected from the group consisting of toxins and anti-retroviral agents.

40. The hetero-multimeric gelsolin fusion construct according to claim 34, said construct comprising a recognition molecule and a reporter group.

41. The hetero-multimeric gelsolin fusion construct according to claim 34, said construct comprising at least two immunogens.

42. The gelsolin fusion construct according to claim 34, said construct comprising a vesicle that consists essentially of PIP or PIP₂.

43. The gelsolin fusion construct according to claim 34, wherein the vesicle comprises lipids selected from the group consisting of PC, PE and PS.

44. The gelsolin fusion construct according to claim 34, said construct comprising a mixed lipid vesicle.

45. The gelsolin fusion construct according to claim 34, wherein the vesicle comprises a detergent.

46. The gelsolin fusion construct according to claim 34, wherein said vesicle contains a bioactive agent.

47. A lipid binding protein fusion construct comprising a lipid binding protein fusion polypeptide and a vesicle comprising a lipid capable of binding to said lipid binding protein fusion polypeptide, said construct being multimeric or hetero-multimeric.

- 84 -

48. The lipid binding protein fusion construct according to claim 47, wherein the lipid binding protein is selected from the group consisting of villin, severin, fragmin, profilin, cofilin, Cap42(a), gCap39, Cap2 and destrin.

49. The lipid binding protein fusion construct according to claim 47, wherein the lipid binding protein is protein kinase C, lipocortin or DNase I.

50. A method for producing a multimeric or hetero-multimeric gelsolin fusion polypeptide comprising the step of transforming a unicellular host with a recombinant DNA molecule comprising a DNA sequence encoding a gelsolin fusion polypeptide operatively linked to an expression control sequence.

51. A method for treating a patient having AIDS, ARC, HIV infection or antibodies to HIV comprising the step of administering to the patient a therapeutically effective amount of a multimeric or hetero-multimeric CD4-gelsolin fusion construct.

52. The method according to claim 51 wherein the fusion construct comprises a toxin or an anti-retroviral agent.

53. A method for identifying the presence of a target molecule in a sample comprising the step of contacting the sample with a hetero-multimeric gelsolin fusion construct according to claim 40.

54. A method for identifying the presence of a target molecule in vivo comprising the step of administering to a patient an effective amount of a

- 85 -

hetero-multimeric gelsolin fusion construct according to
claim 40.

FIG. 1A

1 ATGGCTCCGACCGCCCGCCCGCGCTGCTTTGCGCGCTGTCCCTGGCGCTGTGCGCG 60
 -27 MetAlaProHisArgProAlaProAlaLeuCysAlaLeuSerLeuAlaLeuCysAla

61 CTGTCGCTGCCCGTCCGCGCGGCCACTGCGTCGCGGGGGCGGTCCAGCGGGGGCGCCC 120
 -7 LeuSerLeuProValArgAlaAlaThrAlaSerArgGlyAlaSerGlnAlaGlyAlaPro

121 CAGGGCGGGTGCCCGAGGCGCGGCCCAACAGCATGGTGTTGGAAACACCCCGAGTTCCCTC 180
 14 GlnGlyArgValProGluAlaArgProAsnSerMetValValGluHisProGluPheLeu

181 AAGGCAGGGAAGGAGCCTGGCCTGCAGATCTGGCGTGTGGAGAAGTTCGATCTGGTGCCCC 240
 34 LysAlaGlyLysGluProGlyLeuGlnIleTrpArgValGluLysPheAspLeuValPro

241 GTGCCCCACCAACCTTTATGGAGACTTCTTACGGGCGACGCCCTACGTCACTCCTGAAGACA 300
 54 ValProThrAsnLeuTyrGlyAspPhePheThrGlyAspAlaTyrValIleLeuLysThr

301 GTGCAGCTGAGGAACGGAAATCTGCAGTATGACCTCCACTACTGGCTGGGCAATGAGTGC 360
 74 ValGlnLeuArgAsnGlyAsnLeuGlnTyrAspLeuHisTyrTrpLeuGlyAsnGluCys

361 AGCCAGGATGAGAGCGGGCGGCCCATCTTTACCGTGCAGCTGGATGACTACCTGAAC 420
 94 SerGlnAspGluSerGlyAlaAlaAlaIlePheThrValGlnLeuAspAspTyrLeuAsn

FIG. 1B

421	GGCCGGGCGGTGCAGCACCGGTGAGGTCCAGGGCTTCGAGTCGGCCACCTTCCTAGGCTAC	480
114	GlyArgAlaValGlnHisArgGluValGlnGlyPheGluSerAlaThrPheLeuGlyTyr	
481	TTCAAGTCTGGCCTGAAGTACAAGAAAGAGGTGTGGCATCAGGATTCAAGCACGTGGTA	540
134	PheLysSerGlyLeuLysTyrLysLysGlyGlyValAlaSerGlyPheLysHisValVal	
541	CCCAACGAGGTGGTGCAGAGACTCTTCCAGGTCAAAGGGCGGTGGTCCGTGCC	600
154	ProAsnGluValValGlnArgLeuPheGlnValLysGlyArgArgValValArgAla	
601	ACCGAGGTACCTGTGTCTGGGAGAGCTTCAACAAATGGCGACTGCTTCATCCTGGACCTG	660
174	ThrGluValProValSerTrpGluSerPheAsnAsnGlyAspCysPheIleLeuAspLeu	
661	GGCAACAACATCCACCAGTGGTGTGGTTCCAACAGCAATCGGTATGAAAGACTGAAGGCC	720
194	GlyAsnAsnIleHisGlnTrpCysGlySerAsnSerAsnArgTyrGluArgLeuLysAla.	
721	ACACAGGTGTCCAAGGGCATCCGGGACAAACGAGCGGAGTGGCCGGCCGAGTGCACGTG	780
214	ThrGlnValSerLysGlyIleArgAspAsnGluArgSerGlyArgAlaArgValHisVal	
781	TCTGAGGAGGGCACTGAGCCCGAGGCGATGCTCCAGGTGCTGGGCCCCCAAGCCGGCTCTG	840
234	SerGluGluGlyThrGluProGluAlaMetLeuGlnValLeuGlyProLysProAlaLeu	

2/42

FIG. 1C

841	CCTGCAGGTACCGAGGACACCGCCAGGAGGATGCGGCCAACCGCAAGCTGGCCAAAGCTC	900
254	ProAlaGlyThrGluAspThrAlaLysGluAspAlaAlaAsnArgLysLeuAlaLysLeu	
901	TACAAGGTCTCCAATGGTGCAGGGACCATGTCCGTCTCCCTCGTGGCTGATGAGAAACCCC	960
274	TyrLysValSerAsnGlyAlaGlyThrMetSerValSerLeuValAlaAspGluAsnPro	
961	TTCGCCCCAGGGGCCCTGAAGTCAGAGGACTGCTTCATCCTGGACCACGGCAAAGATGGG	1020
294	PheAlaGlnGlyAlaLeuLysSerGluAspCysPheIleLeuAspHisGlyLysAspGly	
1021	AAAAATCTTTGTCTGGAAAGCAAGCAGGCAAAACACGGAGGAGAGGCTGCCCTCAA	1080
314	LysIlePheValTrpLysGlyLysGlnAlaAsnThrGluGluArgLysAlaAlaLeuLys	
1081	ACAGCCTCTGACTTCATCACCAAGATGGACTACCCCAAGCAGACTCAGGTCTCGGTCCTT	1140
334	ThrAlaSerAspPheIleThrLysMetAspTyrProLysGlnThrGlnValSerValLeu	
1141	CCTGAGGGCGGTGAGACCCCACTGTTCAAGCAGTTCTTCAAGAACTGGCGGACCCAGAC	1200
354	ProGluGlyGlyGluThrProLeuPheLysGlnPhePheLysAsnTrpArgAspProAsp	
1201	CAGACAGATGGCCTGGGCTTGTCCTACCTTTCCAGCCATATCGCCAAACGTGGAGCGGGTG	1260
374	GlnThrAspGlyLeuGlyLeuSerTyrLeuSerSerHisIleAlaAsnValGluArgVal	

3/42

4/42

1261	CCCTTCGACGCCCGCCACCCCTGCACACCTCCACTGCCATGGCCGCCAGCACGGCATGGAT	1320
394	ProPheAspAlaAlaThrLeuHisThrSerThrAlaMetAlaAlaGlnHisGlyMetAsp	
1321	GACGATGGCAGGCCAGAAACAGATCTGGAGAAATCGAAGGTTCCAACAAGGTGCCCGTG	1380
414	AspAspGlyThrGlyGlnLysGlnIleTrpArgIleGluGlySerAsnLysValProVal	
1381	GACCCCTGCCACATATGGACAGTTCTATGGAGGCGACAGCTACATCATTTCTGTACAACTAC	1440
434	AspProAlaThrTyrglyGlnPheTyrglyGlyAspSerTyriIleLeuTyraSnTyri	
1441	CGCCATGGTGGCCCGCCAGGGGCAGATAATCTATAACTGGCAGGGTGCCCCAGTCTACCCAG	1500
454	ArgHisGlyGlyArgGlnGlyGlnIleIleTyraSnTrpGlnGlyAlaGlnSerThrGln	
1501	GATGAGGTCGCTGCATCTGCCATCCTGACTGCTCAGCTGGATGAGGAGCTGGAGGTACC	1560
474	AspGluValAlaAlaSerAlaIleLeuThrAlaGlnLeuAspGluGluLeuGlyGlyThr	
1561	CCTGTCCAGAGCCCGTGTGGTCCAAGGCAAGAGCCCGCCACCTCATGAGCCTGTTGGT	1620
494	ProValGlnSerArgValValGlnGlyLysGluProAlaHisLeuMetSerLeuPheGly	
1621	GGGAAGCCCATGATCATCTACAAGGGCGGCACCTCCCGCAGGGCGGCAGACAGCCCCCT	1680
514	GlyLysProMetIleIleTyriLysGlyGlyThrSerArgGluGlyGlyGlnThrAlaPro	

FIG. 1D

FIG. 1E

1681	GCCAGCACCCGCTCTTCCAGGTCCGCGCAACAGCGCTGGAGCCACCCGGGCTGTTGAG	1740
534	AlaSerThrArgLeuPheGlnValArgAlaAsnSerAlaGlyAlaThrArgAlaValGlu	
1741	GTATTGCCCTAAGGCTGGTGCACTGAACCTCAACGATGCCCTTGTCTGAAAACCCCTCA	1800
554	ValLeuProLysAlaGlyAlaLeuAsnSerAsnAspAlaPheValLeuLysThrProSer	
1801	GCCGCCCTACCTGTGGGTGGGTACAGGAGCCAGCGAGGAGAGAGAACGCGGGCCAGGAG	1860
574	AlaAlaTyrLeuTrpValGlyThrGlyAlaSerGluAlaGluLysThrGlyAlaGlnGlu	
1861	CTGCTCAGGGTGCTCGGGGCCCAACCTGTGCAGGTGGCAGAGGAGGAGCCAGATGGC	1920
594	LeuLeuArgValLeuArgAlaGlnProValGlnValAlaGluGlySerGluProAspGly	
1921	TTCTGGGAGGCCCTGGGGCGGAAGGCTGCCCTACCGCACATCCCCACGGCTGAAGGACAAG	1980
614	PheTrpGluAlaLeuGlyGlyLysAlaAlaTyrArgThrSerProArgLeuLysAspLys	
1981	AAGATGGATGCCCATCCTCCTCGCCCTCTTTGCCCTGCTCCCAACAAGATTGGACGTTTGTG	2040
634	LysMetAspAlaHisProProArgLeuPheAlaCysSerAsnLysIleGlyArgPheVal	
2041	ATCGAAGAGGTTCTTGGTGAGCTCATGCAGGAAGACCTGGCAACGGATGACGTGCTGCTT	2100
654	IleGluGluValProGlyGluLeuMetGlnGluAspLeuAlaThrAspAspValMetLeu	

FIG. 1F

2101 CTGGACACCTGGGACCAAGTCTTTGTCTGGTTGGAAAGGATTCTCAAGAAGAAAAAG 2160
674 LeuAspThrTrpAspGlnValPheValTrpValGlyLysAspSerGlnGluGluLys

2161 ACAGAAAGCCCTTGACTTCTGCTAAGCGGTACATCGAGACGGACCCAGCCAATCGGGATCGG 2220
694 ThrGluAlaLeuThrSerAlaLysArgTyrIleGluThrAspProAlaAsnArgAspArg

2221 CGGACGCCCCATCACCGTGGTGAAGCAAGGCTTTTGAGCCTCCCTCTTGTGGGCTGGTTC 2280
714 ArgThrProIleThrValValLysGlnGlyPheGluProProSerPheValGlyTrpPhe

2281 CTTGGCTGGGATGATGATTACTGGTCTGTGGACCCCTTTGGACAGGGCCCATGGCTGAGCTG 2340
734 LeuGlyTrpAspAspTyrTrpSerValAspProLeuAspArgAlaMetAlaGluLeu

2341 GCTGCCTGAGGAGGGGCAGGGCCCCACCCATGTCAACCGGTCAAGTGCCCTTTTGGAACTGTCC 2400
754 AlaAla*

2401 TTCCCTCAAAGAGGCCTTAGAGCGAGCAGAGCAGCTCTGCTATGAGTGTGTGTGTGTG 2460

2461 TGTGTTGTTCTTTTCTTTTTTTACAGTATCCAAAAATAGCCCTGCAAAAATTCAGAG 2520

2521 TCCTTGCAAAATGTCTAAAATGTCAGTGTGTGGGAAATTAAATCCAATAAAACATTTT 2580

2581 GAAGTGTG 2588

6/42

7/42

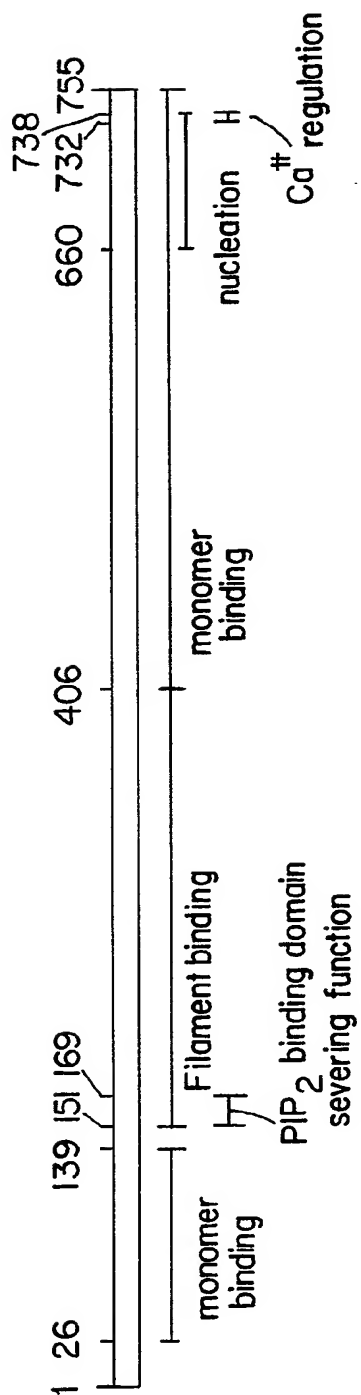


FIG. 2

FIG. 3A

1 ATGAACCGGGAGTCCCTTTTAGGCACTTGCTTCTGGTGCTGCAACTGGCGCTCCTCCCA 60
 - 25 MetAsnArgGlyValPropheArgHisLeuLeuValLeuGlnLeuAlaLeuLeuPro -6

 61 GCAGCCACTCAGGGAAGAAAGTGGTGCTGGGCAAAAAGGGGATACAGTGGAACTGACC 120
 - 5 AlaAlaThrGlnGlyLysLysValValLeuGlyLysLysGlyAspThrValGluLeuThr 15

 121 TGTACAGCTTCCCAGAAAGAGCATAACAATTCCACTGGAAAACTCCAACCATGAAAG 180
 16 CysThrAlaSerGlnLysLysSerIleGlnPheHisTrpLysAsnSerAsnGlnIleLys 35

 181 ATTCTGGGAAATCAGGGCTCCTTCTTAACTAAAGGTCCATCCCAAGCTGAATGATCGCGCT 240
 36 IleLeuGlyAsnGlnGlySerPheLeuThrLysGlyProSerLysLeuAsnAspArgAla 55

 241 GACTCAAGAAGAGCTTGTGGGACCAAGGAACTTCCCTGATCATCAAGAATCTTAAG 300
 56 AspSerArgArgSerLeuTrpAspGlnGlyAsnPheProLeuIleIleLysAsnLeuLys 75

 301 ATAGAAGACTCAGATACATCTGTGAAGTGGAGGACCAGAAAGGAGGAGTGCAATTG 360
 76 IleGluAspSerAspThrTyrIleCysGluValGluAspGlnLysGluGluValGlnLeu 95

FIG. 3B

361	CTAGTGTTCGGATTGACTGCCAACTCTGACACCCACCTGCTTCAGGGGCAGAGCCTGACC	420
96	LeuValPheGlyLeuThrAlaAsnSerAspThrHisLeuLeuGlnGlyGlnSerLeuThr	115
421	CTGACCTTGGAGAGCCCCCTGGTAGTAGCCCCCTCAGTGCAATGTAGGAGTCCAAAGGGGT	480
116	LeuThrLeuGluSerProProGlySerSerProSerValGlnCysArgSerProArgGly	135
481	AAAAACATACAGGGGGGAAGACCCCTCTCCGTGTCTCAGCTGGAGCTCCAGGATAGTGGC	540
136	LysAsnIleGlnGlyGlyLysThrLeuSerValSerGlnLeuGluLeuGlnAspSerGly	155
541	ACCTGGACATGCACGTGTCTTGCAGAACCAAGAGGTGGAGTTCAAAATAGACATCGTG	600
156	ThrTrpThrCysThrValLeuGlnAsnGlnLysLysValGluPheLysIleAspIleVal	175
601	GTGCTAGCTTCCAGAGGCCTCCAGCATAGTCTACAAGAAAGAGGGGGAACAGGTGGAG	660
176	ValLeuAlaPheGlnLysAlaSerSerIleValTyrLysLysGluGlyGluGlnValGlu	195
661	TTCTCCTTCCCACTCGCCTTTACAGTTGAAAAGCTGACGGGCAGTGGCGAGCTGTGGTGG	720
196	PheSerPheProLeuAlaPheThrValGluLysLeuThrGlySerGlyGluLeuTrpTrp	215

FIG. 3C

10/42

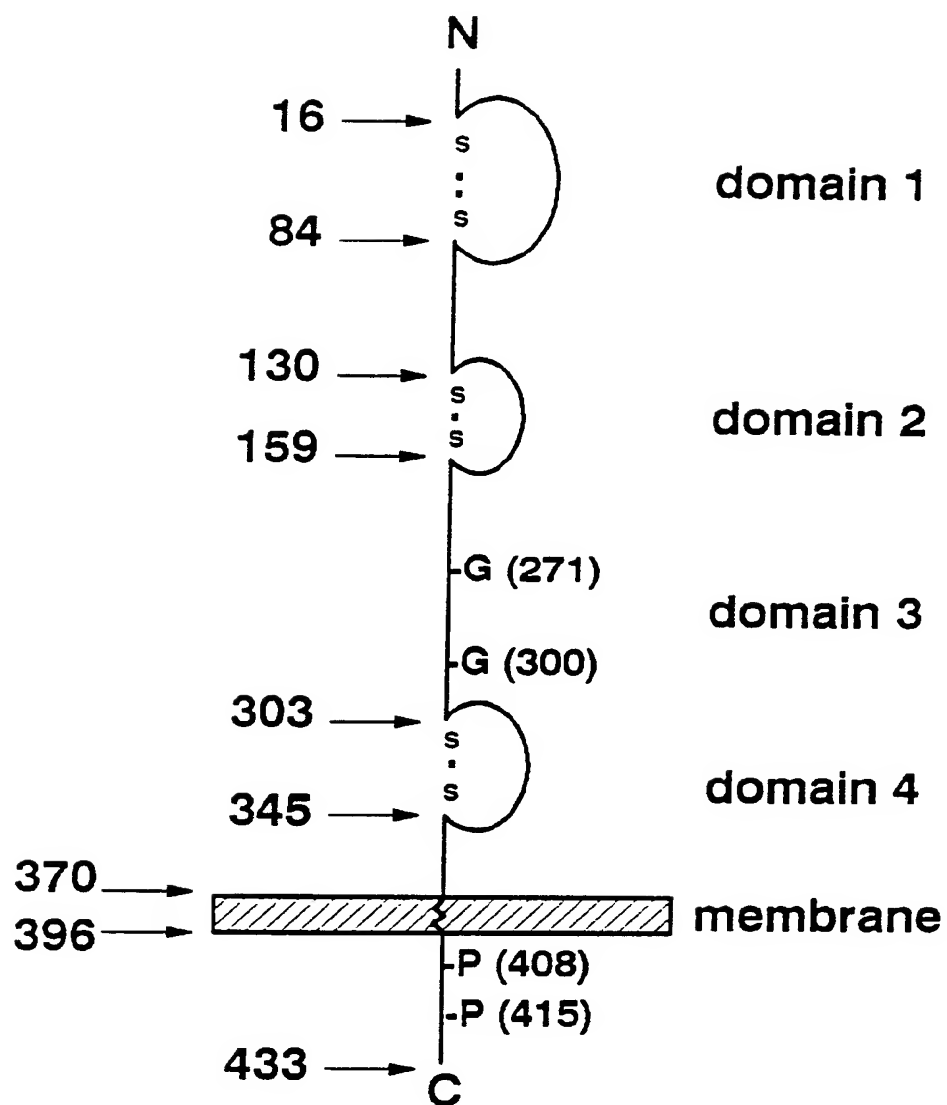
721	CAGCGGAGAGGGCTTCCTCCTCCAAGTCTTGGATCACCTCTGACCTGAAGAACAAGGAA	780
216	GlnAlaGluArgAlaSerSerSerLysSerTrpIleThrSerAspLeuLysAsnLysGlu	235
781	GTGTCTGTAAACGGGTTACCCAGGACCCTAAGCTCCAGATGGGCAAGAAGCTCCCGCTC	840
236	ValSerValLysArgValThrGlnAspProLysLeuGlnMetGlyLysLysLeuProLeu	255
841	CACCTCACCCCTGCCCCAGGCCCTTGCCCTCAGTATGCTGGCTCTGGAAACCTCACCCCTGGCC	900
256	HisLeuThrLeuProGlnAlaLeuProGlnTyrAlaGlySerGlyAsnLeuThrLeuAla	275
901	CTTGAAGCGAAACAGGAAAGTTGCATCAGGAAGTGAACCTGGTGGTGATGAGAGCCACT	960
276	LeuGluAlaLysThrGlyLysLeuHisGlnGluValAsnLeuValMetArgAlaThr	295
961	CAGCTCCAGAAAAATTTGACCTGTGAGGTGTGGGGACCCACCTCCCTAAGCTGATGCTG	1020
296	GlnLeuGlnLysAsnLeuThrCysGluValTrpGlyProThrSerProLysLeuMetLeu	315
1021	AGCTTGAAACTGGAGAACAAAGGAGGCAAGGTCTCGAAGCGGGAGAAGCGGTGTGGGTG	1080
316	SerLeuLysLeuGluAsnLysGluAlaLysValSerLysArgGluLysAlaValTrpVal	335

FIG. 3D

1081	CTGAACCTGAGCGGGGATGTGGCAGTGTCTGCTGAGTGACTCGGGACAGGTCCTGCTG	1140
336	LeuAsnProGluAlaGlyMetTrpGlnCysLeuLeuSerAspSerGlyGlnValLeuLeu	355
1141	GAATCCAACATCAAGGTTCTGTGCCCACATGGTCCACCCCGGTGCAGCCAATGGCCCTGATT	1200
356	GluSerAsnIleLysValLeuProThrTrpSerThrProValGlnProMetAlaLeuIle	375
1201	GTGCTGGGGGCGTCGCGCCCTCCTGCTTTTCATTGGGCTAGGCATCTTCTCTGTGTC	1260
376	ValLeuGlyGlyValAlaGlyLeuLeuLeuPheIleGlyLeuGlyIlePhePheCysVal	395
1261	AGGTGCCGGCACCGAAGGCGCCCAAGCAGAGCGGATGTCTCAGATCAAGAGACTCCTCAGT	1320
396	ArgCysArgHisArgArgArgGlnAlaGluArgMetSerGlnIleLysArgLeuLeuSer	415
1321	GAGAAGAAGACCTGCCAGTGCCCTCACCGGTTTCAGAAGACATGTAGCCCATTTGA	1377
416	GluLysLysThrCysGlnCysProHisArgPheGlnLysThrCysSerProIleEnd	434

11/42

12/42

CD4**FIG. 4**

14/42

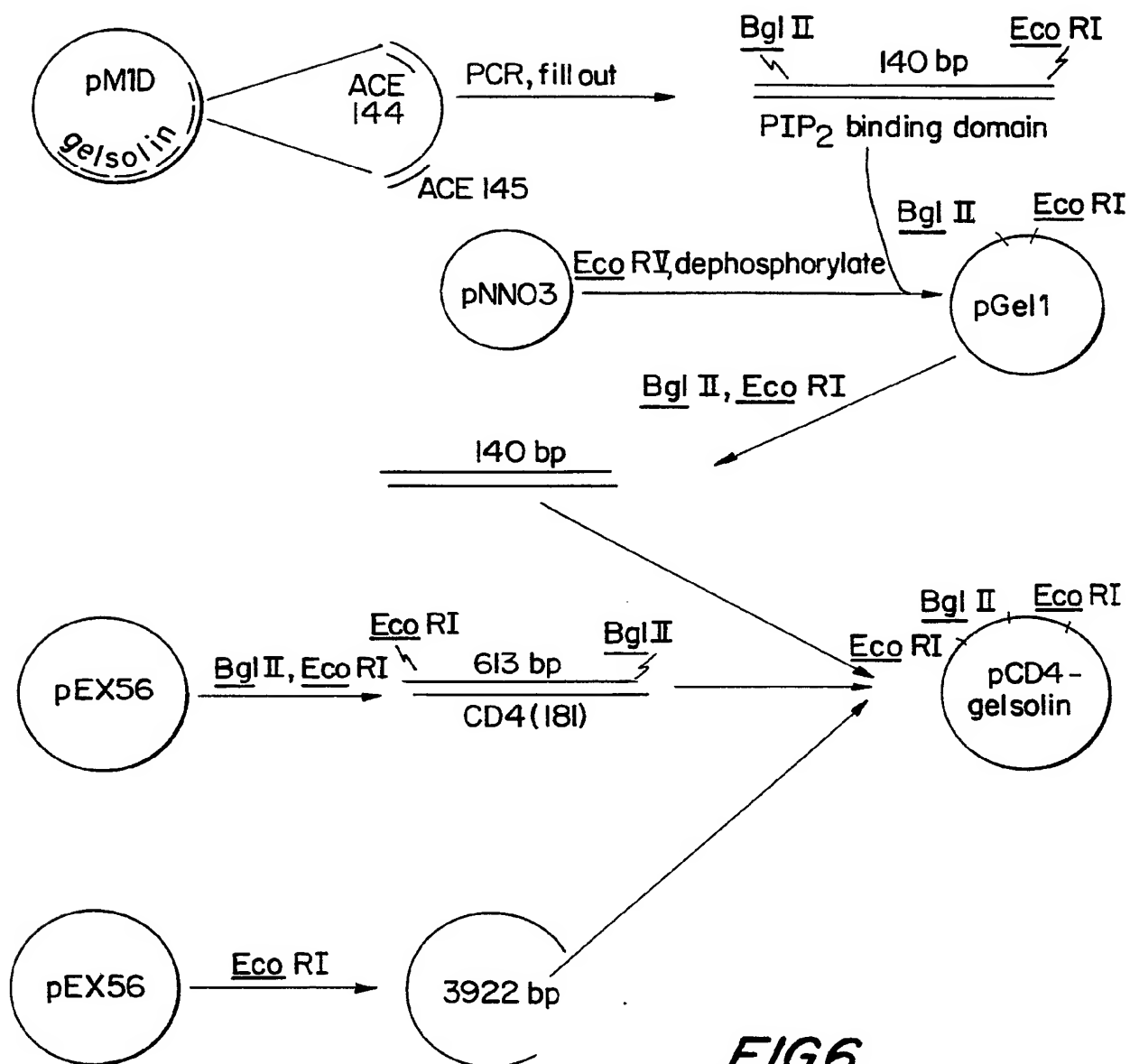


FIG. 7A

1 ATGAAAAAGTAGTACTGGGCAAAAAGGGGATACAGTGGAACCTGACCTGTACAGCTTCC 60
 1 MetLysLysValValLeuGlyLysLysGlyAspThrValGluLeuThrCysThrAlaSer 20

 61 CAGAAGAAGAGCATACAAATTCCTGGAATACTCCAACCCAGATAAAGATTCTGGGAAAT 120
 21 GlnLysLysSerIleGlnPheHisTrpLysAsnSerAsnGlnIleLysIleLeuGlyAsn 40

 121 CAGGGCTCCTTCTTAACATAAAGGTCCATCCAAGCTGAATGATCGCGCTGACTCAAGAAGA 180
 41 GlnGlySerPheLeuThrLysGlyProSerLysLeuAsnAspArgAlaAspSerArgArg 60
 15/42

 181 AGCTTGTGGGACCAAGGAACTTTCCCTGATCATCAAGAAATCTTAAGATAGAAAGACTCA 240
 61 SerLeuTrpAspGlnGlyAsnPheProLeuIleIleLysAsnLeuLysIleGluAspSer 80

 241 GATACTTACATCTGTGAAGTGGAGGACCAGAAAGAAAGATTTCAGCTGCTGGTTTTCGGA 300
 81 AspThrTyriIleCysGluValGluAspGlnLysGluGluValGlnLeuLeuValPheGly 100

 301 TTGACTGCCAACTCTGACACCCACCTGCTTCAGGGGCAGAGCCTGACCCCTTGAG 360
 101 LeuThrAlaAsnSerAspThrHisLeuLeuGlnGlyGlnSerLeuThrLeuThrLeuGlu 120

FIG. 7B

361	AGCCCCCTGGTAGTAGCCCTCAGTGCAATGTAGGAGTCCAAGGGGTAAACATACAG	420
121	SerProGlySerSerProSerValGlnCysArgSerProArgGlyLysAsnIleGln	140
421	GGGGGAAGACCCCTCTCCGTGTCTCAGCTGGAGCTCCAGGATAGTGGCACCTGGACATGC	480
141	GlyGlyLysThrLeuSerValSerGlnLeuGluLeuGlnAspSerGlyThrTrpThrCys	160
481	ACTGTCTTGCAGAACCCAGAAAGGTGGAGTTCAAAAATAGACATCGTGGTGCTAGCTTTC	540
161	ThrValLeuGlnAsnGlnLysLysValGluPheLysIleAspIleValValLeuAlaPhe	180
541	CAGAAGGGGAAGATCTACGGGGCGGTGGCATCAGGATTCAGCACGTGGTACCCAACGAG	600
181	GlnLysGlyLysIleTyrGlyGlyValAlaSerGlyPheLysHisValValProAsnGlu	200
	-----spacer----- -----gelsolin	
601	GTGGTGGTGACAGACTCTTCCAGGTCAAGGGCGGTGGTGGTCCGTGCCCTAA	
201	ValValGlnArgLeuPheGlnValLysGlyArgArgValValArgAla*	
	amino acids 150-173-----	

16/42

17/42

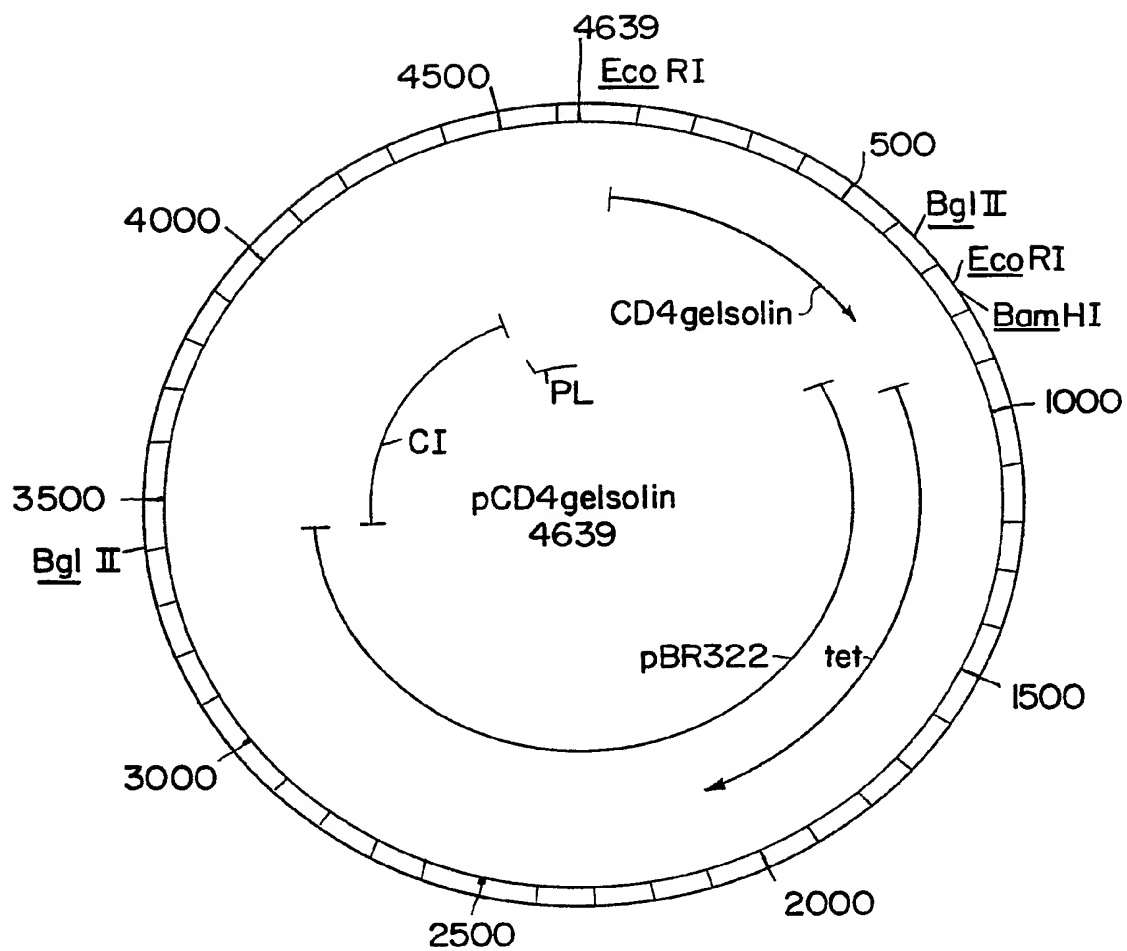
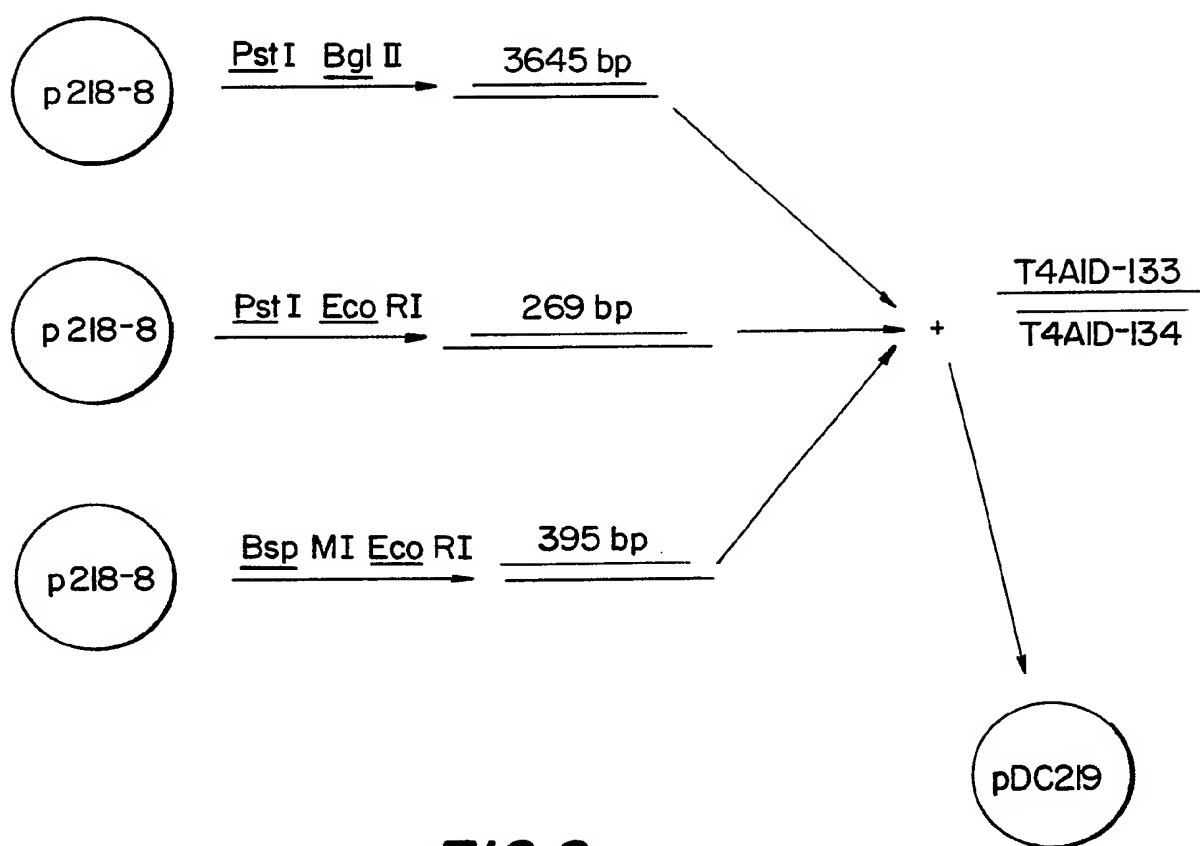


FIG. 8

18/42

**FIG. 9**

19/42

FIG. 10A

1 GAATTCTTAC ACTTAGTTAA ATTGCTAACT TTATAGATTA CAAAACTTAG
51 GAAATCGATT TGGATGAAAA AAGTAGTACT GGGCAAAAAA GGGGATACAG
101 TGGAAC TGAC CTGTACAGCT TCCCAGAAGA AGAGCATACA ATTCCACTGG
151 AAAA ACTCCA ACCAGATAAA GATTCTGGGA AATCAGGGCT CCTTCTTAAC
201 TAAAGGTTCCA TCCAAGCTGA ATGATCGCGC TGACTCAAGA AGAAGCTTGT
251 GGGACCAAGG AAAC TTCC CTGATCATCA AGAATCTTAA GATAGAAGAC
301 TCAGATACTT ACATCTGTGA AGTGGAGGAC CAGAAGGAGG AGGTGCAATT
351 GCTAGTGTTT GGA TTGACTG CCAACTCTGA CACCCACCTG CTCAGGGGT
401 GATAGTAAGA TCCTGCAGCC CAGCTTGGGG ACCCTAGAGG TCCCCTTTT
451 TATTTTGAAT TGGGAGATCC CAATTCTCAT GTTTGACAGC TTATCATCGA
501 TAAGCTAGCT TTAATGCGGT AGTTTATCAC AGTTAAATTG CTAACGCAGT
551 CAGGCACCGT GTATGAAATC TAACAATGCG CTCATCGTCA TCCTCGGCAC
601 CGTCACCCCTG GATGCTGTAG GCATAGGCTT GGTATGCCG G TACTGCCCG
651 GCCTCTTGCG GGATATCGTC CATTCCGACA GCATCGCCAG TCACTATGGC
701 GTGCTGCTAG CGCTATATGC GTTGATGCAA TTTCTATGCG CACCCGTTCT

20/42

FIG. 10B

751 CGGAGCACTG TCCGACCGCT TTGGCCGCCG CCCAGTCCTG CTCGCTTCGC
801 TACTTGGAGC CACTATCGAC TACGCGATCA TGGCGACCAC ACCCGTCCTG
851 TGGATTCTCT ACGCCGGACG CATCGTGGCC GGCATCACCG GCGCCACAGG
901 TCGGGTTGCT GCGCCCTATA TCGCCGACAT CACCGATGGG GAAGATCGGG
951 CTCGCCACTT CGGGCTCATG AGCGCTTGTT TCGGCGTGGG TATGGTGGCA
1001 GGCCCCGTGG CCGGGGGACT GTTGGGGGCC ATCTCCTTGC ACGCACCATT
1051 CCTTGGGGCG GCGGTGCTCA ACGGCCCTCA CCTACTACTG GGCTGCTTCC
1101 TAATGCAGGA GTCGCATAAG GGAGAGCGTC GTCCGATGCC CTTGAGAGCC
1151 TTCAACCCAG TCAGCTCCTT CCGGTGGGCG CGGGGCATGA CTATCGTTCG
1201 CGCACTTATG ACTGTCTTCT TTATCATGCA ACTCGTAGGA CAGGTGCCGG
1251 CAGCGCTCTG GGTCAATTTT CCGGAGGACC GCTTTCGCTG GAGCGCGACG
1301 ATGATCGGCC TGTGCGTTGC GGTATTCGGA ATCTTGCACG CCCTCGCTCA
1351 AGCCTTCGTC ACTGGTCCCG CCACCAAACG TTTCGGCGAG AAGCAGGCCA
1401 TTATCGCCCG CATGGCGGCC GACGCGCTGG GCTACGTCTT GCTGGCGTTC
1451 GCGACGCGAG GCTGGATGGC CTTCCCCATT ATGATTCTC TCGCTTCCGG

21/42

FIG. 10C

1501 CGGCATCGGG ATGCCCCGGT TGCAGGCCAT GCTGTCCAGG CAGGTAGATG
1551 ACGACCATCA GGGACAGCTT CAAGGATCGC TCGCGGCTCT TACCAGCCTA
1601 ACTTCGATCA CTGGACCGCT GATCGTCAGG GCGATTTATG CCGCCTCGGC
1651 GAGCACATGG AACGGGTTGG CATGGATTGT AGGCGCCGCC CTATACCTTG
1701 TCTGCCTCCC CGCGTTGCGT CGCGGTGCAT GGAGCCGGGC CACCTCGACC
1751 TGAATGGAAG CCGGCGGCAC CTCGCTAAGG GATTCACCAC TCCAAGAATT
1801 GGAGCCAATC AATTCTTGCG GAGAACTGTG AATGCGCAA CCAACCCCTG
1851 GCAGAACATA TCCATCGCGT CCGCCATCTC CAGCAGCCGC ACGCGGCGCA
1901 TCTCGGGGGA TGATCAGCTG CCTCGCGCGT TTCGGTGATG ACGGTGA AAA
1951 CCTCTGACAC ATGCAGCTCC CGGAGACGGT CACAGCTTGT CTGTAAGCGG
2001 ATGCCGGGAG CAGACAAGCC CGTCAGGGCG CGTCAGCGGG TGTGGCGGG
2051 TGTCGGGGCG CAGCCATGAC CCAGTCACGT AGCGATAGCG GAGTGTATAC
2101 TGGCTTAACT ATGCGGCATC AGAGCAGATT GTACTGAGAG TGCACCATAT
2151 GCGGTGTGAA ATACCGCACA GATGCGTAAG GAGAAAATAC CGCATCAGGC
2201 GCTCTTCCGC TTCTCTCGTC ACTGACTCGC TCGGCTCGGT CGTTCGGCTG

22/42

FIG. 10D

2251 CCGCGAGCGG TATCAGCTCA CTCAAAGGCG GTAATACGGT TATCCACAGA
2301 ATCAGGGGAT AACGCAGGAA AGAACATGTG AGCAAAAGGC CAGCAAAAGG
2351 CCAGGAACCG TAAAAAGGCC GCGTTGCTGG CGTTTTTCCA TAGGCTCCGC
2401 CCCCCTGACG AGCATCACAA AAATCGACGC TCAAGTCAGA GGTGGCGAAA
2451 CCCGACAGGA CTATAAAGAT ACCAGGCGTT TCCCCCTGGA AGCTCCCTCG
2501 TCGGCTCTCC TGTTCGGACC CTGCCGCTTA CCGGATACCT GTCCGCCCTT
2551 CTCCCTTCGG GAAGCGTGGC GCTTCTCTCA TGCTCACGCT GTAGGTATCT
2601 CAGTTCGGTG TAGGTCGTTC GCTCCAAGCT GGGCTGTGTG CACGAACCCC
2651 CCGTTCAGCC CGACCGCTGC GCCTTATCCG GTAACATCG TCTTGAGTCC
2701 AACC CGGTAA GACACGACTT ATCGCCACTG GCAGCAGCCA CTGGTAACAG
2751 GATTAGCAGA GCGAGGTATG TAGGCGGTGC TACAGAGTTC TTGAAGTGGT
2801 GGCCTAACTA CCGCTACACT AGAAGGACAG TATTGGTAT CTGCGCTCTG
2851 CTGAAGCCAG TTACCTTTCGG AAAAAGAGTT GGTAGCTCTT GATCCGGCAA
2901 ACAAAACCACC GCTGGTAGCG GTGGTTTTT TGTTCGAAG CAGCAGATTA
2951 CGCGCAGAAA AAAAGGATCT CAAGAAGATC CTTTGATCTT TTCTACGGGG

23/42

FIG. 10E

3001 TCTGACGCTC AGTGGAACGA AACTCACGT TAAGGGATT TGGTCATGAG
3051 ATTATCAAAA AGGATCTTCA CCTAGATCCT TTTTCAGATCT CCCGATCTTT
3101 AGCTGTCTTG GTTTGCCCAA AGCGCATGTC ATAATCTTTC AGGGTTATGC
3151 GTTGTTCCAT ACAACCTCCT TAGTACATGC AACCATTATC ACCGCCAGAG
3201 GTAAAATAGT CAACACGCAC GGTGTTAGAT ATTTATCCCT TCGGGTGATA
3251 GATTTAACGT ATGAGCACAA AAAAGAAACC ATTAACACAA GAGCAGCTTG
3301 AGGACGCAG TCGCCTTAA GCAATTATG AAAAAAGAA AAATGAACTT
3351 GGCTTATCCC AGGAATCTGT CGCAGACAAG ATGGGGATGG GGCAGTCAGG
3401 CGTTGGTGCT TTATTTAATG GCATCAATGC ATTAAATGCT TATAACGCCG
3451 CATTGCTTAC AAAAAATTCTC AAAGTTAGCG TTGAAGAAAT TAGCCCTTCA
3501 ATCGCCAGAG AAATCTACGA GATGTATGAA GCGGTTAGTA TGCAGCCGTC
3551 ACTTAGAAGT GAGTATGAGT ACCCTGTTTT TTCTCATGTT CAGGCAGGGA
3601 TGTTCTCACC TAAGCTTAGA ACCTTTACCA AAGGTGATGC GGAGAGATGG
3651 GTAAGCACAA CCAAAAAAGC CAGTGATTCT GCATTCTGGC TTGAGGTTGA
3701 AGGTAATTCC ATGACCGCAC CAACAGGCTC CAAGCCAAGC TTTCCTGACG

FIG. 10F

24/42

3751 GAATGTTAAT TCTCGTTGAC CCTGAGCAGG CTGTTGAGCC AGGTGATTTC
3801 TGCATAGCCA GACTTGGGGG TGATGAGTTT ACCTTCAAGA AACTAATTAG
3851 GGATAGCGGT CAGGTGTTTT TACAACCACT AAACCCACAG TACCCAAATGA
3901 TCCCATGCAA TGAGAGTTGT TCCGTTGTGG GAAAGTTAT CGCTAGTCAG
3951 TGGCCTGAAG AGACGTTTGG CTGATCGGCA AGGTGTTCTG GTCGGCGCAT
4001 AGCTGATAAC AATTGAGCAA GAATCTTCAT CGGGGCTGCA GCCCACCATG
4051 CGTCCGGCGT AGAGGATCTC TCACCTACCA AACAAATGCCC CCTTGCAAAA
4101 AATAAATTCA TATAAAAAAC ATACAGATAA CCATCTGCGG TGATAAATTA
4151 TCTCTGGCGG TGTTGACATA AATACCACTG GCGGTGATAC TGAGCACATC
4201 AGCAGGACGC ACTGACCACC ATGAAGGTGA CGCTCTTAAA ATTAAGCCCT
4251 GAAGAAGGC AGCATTCAAA GCAGAAGGCT TTGGGGTGTG TGATACGAAA
4301 CGAAGCATT

25/42

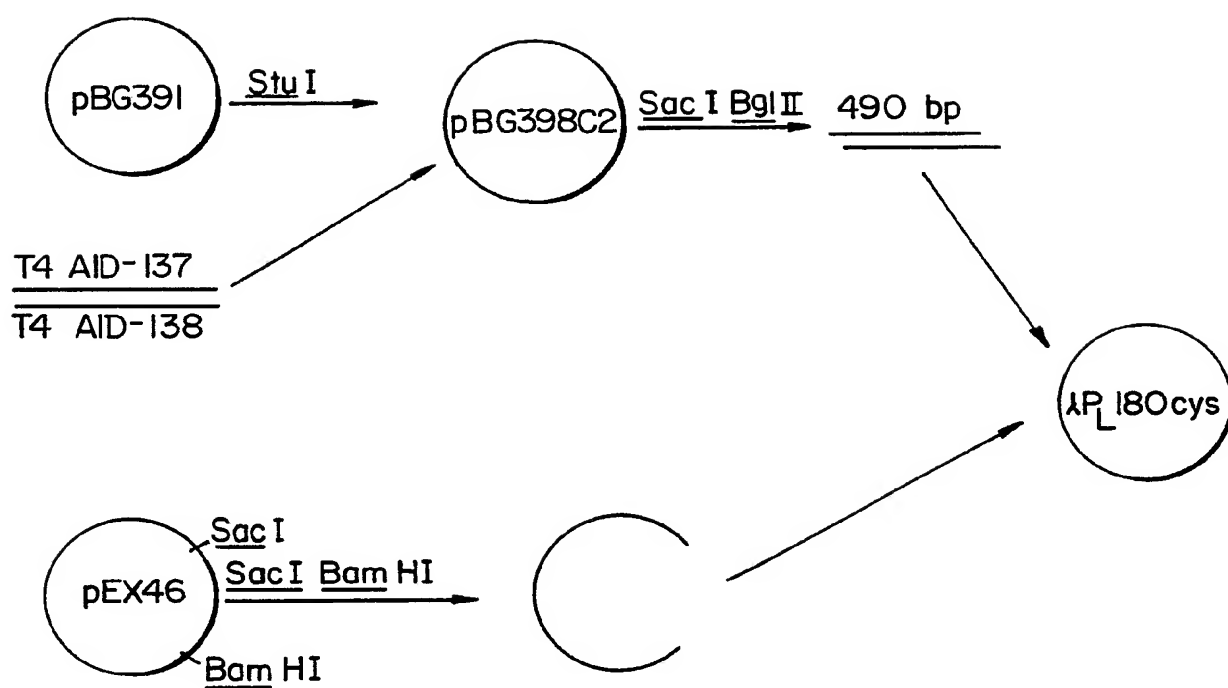


FIG. II

26/42

FIG. 12A

1 GAATTAAATC CAGCTTGCTG TGGAATGTGT GTCAGTTAGG GTGTGGAAG
51 TCCCCAGGCT CCCCAGCAGG CAGAAGTATG CAAAGCATGC ATCTCAATTA
101 GTCAGCAACC AGGTGTGGAA AGTCCCCAGG CTCCCCAGCA GGCAGAAAGTA
151 TGCAAAGCAT GCATCTCAAT TAGTCAGCAA CCATAGTCCC GCCCCTAACT
201 CCGCCCATCC CGCCCCTAAC TCCGCCCAGT TCCGCCCATT CTCCGCCCCCA
251 TGGCTGACTA ATTTTTTTTA TTTATGCAGA GGCCGAGGCC GCCTCGGCCT
301 CTGAGCTATT CCAGAAAGTAG TGAGGAGGCT TTTTGGAGG GGTCCCTCCTC
351 GTATAGAAAC TCGGACCACCT CTGAGACGAA GGCTCGCGTC CAGGCCAGCA
401 CGAAGGAGGC TAAGTGGGAG GGTAGCCGT CGTTGTCCAC TAGGGGTCC
451 ACTCGCTCCA GGGTGTGAAG ACACATGTCG CCCTCTTCGG CATCAAGGAA
501 GGTGATTGGT TTATAGGTGT AGGCCACGTG ACCGGGTGTT CCTGAAGGGG
551 GGCTATAAAA GGGGTGGGG GCGCGTTCGT CCTCACTCTC TTCCGCATCG
601 CTGTCTGCGA GGGCCAGCTG TTGGGCTCGC GGTGAGGAC AAACCTCTCG
651 CGGTCTTTCC AGTACTCTTG GATCGGAAAC CCGTCGGCCT CCGAACGGTA

FIG. 12B

27/42

701 CTCCGCCACC GAGGGACCTG AGCGAGTCCG CATCGACCGG ATCGGAAAC
751 CTCTCGAGAA AGGCGTCTAA CCAGTCACAG TCGCAAGGTA GGCTGAGCAC
801 CGTGGCGGGC GGCAGCGGGT GCGGTCGGG GTTGTCTCTG GCGGAGGTGC
851 TGCTGATGAT GTAATTAAAG TAGGCGGTCT TGAGACGGCG GATGGTCGAG
901 GTGAGGTGTG GCAGGCTTGA GATCGATCTG GCCATACACT TGAGTGACAA
951 TGACATCCAC TTTGCGCTTTC TCTCCACAGG TGTCCACTCC CAGGTCCAAC
1001 TGGATCCAAG CTTGCGACTCG AGGAATTCCC CGAAGGAACA AAGCACCCCTC
1051 CCCACTGGGC TCCTGGTTGC AGAGCTCCAA GTCCTCACAC AGATACGCCT
1101 GTTTGAGAAG CAGCGGGCAA GAAAGACGCA AGCCCAGAGG CCCTGCCATT
1151 TCTGTGGGCT CAGGTCCCTA CTGGCTCAGG CCCCTGCCTC CCTCGGCAAG
1201 GCCACAATGA ACCGGGGAGT CCCTTTTAGG CACTTGCTTC TGGTGCTGCA
1251 ACTGGCGCTC CTCCCAGCAG CCACTCAGG AAAGAAAGTG GTGCTGGGCA
1301 AAAAAGGGGA TACAGTGGA CTGACCTGTA CAGCTTCCCA GAAGAAGAGC
1351 ATACAATTCC ACTGGAAAAA CTCCAACCAG ATAAAGATTG TGGGAAATCA
1401 GGGCTCCTTC TTAACATAAG GTCCATCCAA GCTGAATGAT CGCGCTGACT

28/42

FIG. 12C

1451 CAAGAAGAAG CTTGTGGGAC CAAGGAAACT TTCCCCCTGAT CATCAAGAAT
1501 CTTAAGATAG AAGACTCAGA TACTTACATC TGTGAAGTGG AGGACCAGAA
1551 GGAGGAGGTG CAATTGCTAG TGTTGGGATT GACTGCCAAC TCTGACACCC
1601 ACCTGCTTCA GGGGCAGAGC CTGACCCCTGA CCTTGGAGAG CCCCCCTGGT
1651 AGTAGCCCCCT CAGTGCAATG TAGGAGTCCA AGGGGTAAAA ACATACAGGG
1701 GGGGAAGACC CTCTCCGTGT CTCAGCTGGA GCTCCAGGAT AGTGGCACCT
1751 GGACATGCAC TGCTTTGCAG AACCAGAAGA AGGTGGAGTT CAAAATAGAC
1801 ATCGTGGTGC TAGCTTTCCA GAAGGCCTCC AGCATAGTCT ATAAGAAAAGA
1851 GGGGGAACAG GTGGAGTTCT CCTTCCCCTT CGCCTTTACA GTTGAAAAGC
1901 TGACGGGCAG TGGCGAGCTG TGGTGGCAGG CGGAGAGGGC TTCCCTCCTCC
1951 AAGTCTTGGA TCACCTTTGA CCTGAAGAAC AAGGAAGTGT CTGTAAAACG
2001 GGTTACCCAG GACCCCTAAGC TCCAGATGGG CAAGAAGCTC CCGCTCCACC
2051 TCACCCCTGCC CCAGGCCTTG CCTCAGTATG CTGGCTCTGG AAACCTCACC
2101 CTGGCCCTTG AAGCGAAAAC AGGAAAGTTG CATCAGGAAG TGAACCTGGT

FIG. 12D

2151 GGTGATGAGA GCCACTCAGC TCCAGAAAAA TTTGACCTGT GAGGTGTGGG
2201 GACCCACCTC CCCTAAGCTG ATGCTGAGTT TGAAACTGGA GAACAAGGAG
2251 GCAAAGGTCT CGAAGCGGGA GAAGGCGGTG TGGGTGCTGA ACCCTGAGGC
2301 GGGGATGTGG CAGTGTCTGC TGAGTGACTC GGGACAGGTC CTGCTGGAAT
2351 CCAACATCAA GGTCTTGCCC ACATGGTCGA CCCCGGTGCA GCCAATGGCC
2401 CTGATTTGAG ATCTTTGTGA AGGAACCTTA CTTCTGTGGT GTGACATAAT
2451 TGGACAAACT ACCTACAGAG ATTTAAAGCT CTAAGGTAAA TATAAAATTT
2501 TTAAGTGTAT AATGTGTAA ACTACTGATT CTAATTGTTT GTGTATTTA
2551 GATTCCAACC TATGGAACCTG ATGAATGGA GCAGTGGTGG AATGCCCTTA
2601 ATGAGGAAA CCGTTTTGC TCAGAAGAAA TGCCATCTAG TGATGATGAG
2651 GCTACTGCTG ACTCTCAACA TTCTACTCCT CCAAAAAAGA AGAGAAAGGT
2701 AGAAGACCCC AAGGACTTTC CTTCAGAATT GCTAAGTTTT TTGAGTCATG
2751 CTGTGTTTAG TAATAGAACT CTTGCTTGCT TTGCTATTTA CACCACAAAG
2801 GAAAAAGCTG CACTGCTATA CAAGAAAATT ATGGAAAAAT ATTCTGTAAAC

30/42

FIG. 12E

2851 CTTTATAAGT AGGCATAACA GTTATAATCA TAACATACTG TTTTTTCTTA
2901 CTCACACAG GCATAGAGTG TCTGCTATTA ATAACATATGC TCAAAAATTG
2951 TGTACCTTTA GCTTTTAAAT TTGTAAAGGG GTTAATAAGG AATATTGAT
3001 GTATAGTGCC TTGACTAGAG ATCATAATCA GCCATACCAC ATTTGTAGAG
3051 GTTTACTTG CTTTAAAAA CCTCCCACAC CTCCCCCTGA ACCTGAAACA
3101 TAAAAATGAAT GCAATTGTTG TTGTTAACTT GTTTATTGCA GCTTATAATG
3151 GTTACAAATA AAGCAATAGC ATCACAAATT TCACAAAATA AGCATTTTTT
3201 TCACTGCATT CTAGTTGTGG TTTGTCCAAA CTCATCAATG TATCTTATCA
3251 TGTCTGGATC CTCACGCCG GACGCATCGT GGCCGGCATC ACCGGCGCCA
3301 CAGGTGCGGT TGCTGGCGCC TATATCGCCG ACATCACCGA TGGGGAAGAT
3351 CGGGCTCGCC ACTTCGGGCT CATGAGCGCT TGTTTCGGCG TGGGTATGGT
3401 GGCAGGCCCG TGGCCGGGGG ACTGTGGGC GCCATCTCCT TGCATGCACC
3451 ATTCCTTGGC GCGCGGGTGC TCAACGGCCT CAACCTACTA CTGGGCTGCT
3501 TCCTAATGCA GGAGTCGCAT AAGGGAGAGC GTCGACCGAT GCCCTTGAGA

FIG. 12F

3551 GCCTTCAACC CAGTCAGCTC CTTCCGGTGG GCGCGGGCA TGA CTATCGT
3601 CGCCGCACTT ATGACTGTCT TCTTTATCAT GCAACTCGTA GGACAGGTGC
3651 CGGCAGCGCT CTGGGTCAAT TTCCGGCAGG ACCGCTTTCG CTGGAGCGCG
3701 ACGATGATCG GCCTGTGCGT TCGGGTATTC GGAATCTTGC ACGCCCTCGC
3751 TCAAGCCTTC GTCAC TGGTC CCGCCACCAA ACGTTTCGGC GAGAAGCAGG
3801 CCATTATCGC CGGCATGGCG GCCGACGCG TGGGCTACGT CTTGCTGGCG
3851 TTCGCGACGC GAGGCTGGAT GGCCCTCCCC ATTATGATTC TTCCTCGCTTC
3901 CGGCGGCATC GGGATGCCCC CGTTGCAGGC CATGCTGTCC AGGCAGGTAG
3951 ATGACGACCA TCAGGGACAG CTTCAAGGAT CGCTCGCGGC TCTTACCAGC
4001 CTAAC TCGA TCACTGGACC GCTGATCGTC ACGGCGATT ATGCCGCCCTC
4051 GCGGAGCACA TGGAACGGGT TGGCATGGAT TGTAGGCGCC GCCCTATACC
4101 TTGTCTGCCT CCCCGCGTTG CGTCGCGGTG CATGGAGCCG GGCCACCTCG
4151 ACCTGAATGG AAGCCGGCGG CACCTCGCTA ACGGATTCAC CACTCCAAGA
4201 ATTGGAGCCA ATCAATTCTT GCGGAGAACT GTGAATGCGC AAACCAACCC

32/42

FIG. 12G

4251 TTGGCAGAAC ATATCCATCG CGTCCGCCAT CTCCAGCAGC CGCACGGGGC
4301 GCATCTCGGG CCGCGTTGCT GCGGTTTTTC CATAGGCTCC GCGGGCCCTGA
4351 CGAGCATCAC AAAAATCGAC GCTCAAGTCA GAGGTGGCGA AACCCGACAG
4401 GACTATAAAG ATACCAGGCG TTTCCCCCTG GAAGCTCCCT CGTGCGCTCT
4451 CCTGTTCCGA CCTGCGCGCT TACCGGATAC CTGTCCGCCT TTCTCCCTTC
4501 GGAAGCGTG GCGCTTTCTC AATGCTCAG CTGTAGGTAT CTCAGTTCGG
4551 TGTAGGTCGT TCGCTCCAAG CTGGGCTGTG TGCACGAACC CCCC GTTCAG
4601 CCCGACCGCT GCGCCTTATC CGGTAACAT CGTCTTGAGT CCAACCCGGT
4651 AAGACACGAC TTATCGCCAC TGGCAGCAGC CACTGGTAAC AGGATTAGCA
4701 GAGCGAGGTA TGTAGGCGGT GCTACAGAGT TCTTGAAGTG GTGGCCCTAAC
4751 TACGGCTACA CTAGAAGGAC AGTATTGGT ATCTGCGCTC TGCTGAAGCC
4801 AGTTACCTTC GGA AAAAGAG TTGGTAGCTC TTGATCCGGC AAACAAACCA
4851 CCGCTGGTAG CCGTGGTTTT TTTGTTTGCA AGCAGCAGAT TACGCGCAGA
4901 AAAAAAGGAT CTCAGAAGA TCCTTTGATC TTTTCTACGG GGTCTGACGC

33/42

FIG. 12H

4951 TCAGTGAAC GAAAACTCAC GTTAAGGAT TTTGGTCATG AGATTATCAA
5001 AAAGGATCTT CACCTAGATC CTTTAAATT AAAAATGAAG TTTTAAATCA
5051 ATCTAAAGTA TATATGAGTA AACTTGGTCT GACAGTTACC AATGCTTAAT
5101 CAGTGAGGCA CCTATCTCAG CGATCTGTCT ATTTCGTTCA TCCATAGTTG
5151 CCTGACTCCC CGTCGTGTAG ATAAC TACGA TACGGGAGGG CTTACCATCT
5201 GGGCCCCAGTG CTGCAATGAT ACCGCGAGAC CCACGCTCAC CGGCTCCAGA
5251 TTTATCAGCA ATAAACCAGC CAGCCGGAAG GGCCGAGCGC AGAAGTGGTC
5301 CTGCAACTTT ATCCGCCCTCC ATCCAGTCTA TTAATTGTTG CCGGGAAGCT
5351 AGAGTAAGTA GTTCGCCAGT TAAAGTTTG CGCAACGTTG TTGCCATTGC
5401 TGCAGGCATC GTGGTGCAC GCTCGTCGTT TGGTATGGCT TCATTCAGCT
5451 CCGGTTCCCA ACGATCAAGG CGAGTTACAT GATCCCCCAT GTTGTGCAAA
5501 AAAGCGGTTA GCTCCTTCGG TCCTCCGATC GTTGTCAGAA GTAAGTTGGC
5551 CGCAGTGTTA TCACTCATGG TTATGGCAGC ACTGCATAAT TCTCTTACTG
5601 TCATGCCCATC CGTAAGATGC TTTTCTGTGA CTGGTGAGTA CTCAACCAAG

34/42

FIG. 12I

5651 TCATTCTGAG AATAGTGTAT GCGGCGACCG AGTTGCTCTT GCCCGGCGTC
5701 AACACGGGAT AATACCGCGC CACATAGCAG AACTTTAAAA GTGCTCATCA
5751 TTGGAAAACG TTCTTCGGGG CGAAAACCTCT CAAGGATCTT ACCGCTGTTG
5801 AGATCCAGTT CGATGTAACC CACTCGTGCA CCCAACTGAT CTCAGCATC
5851 TTTTACTTTC ACCAGCGTTT CTGGGTGAGC AAAAACAGGA AGCAAAATG
5901 CCGCAAAAAA GGGAATAAGG GCGACACGGA AATGTTGAAT ACTCATCTC
5951 TTCCCTTTTC AATATTATTG AAGCATTAT CAGGGTTATT GTCTCATGAG
6001 CGGATACATA TTTGAATGTA TTAGAAAAA TAAACAAATA GGGTTCCGC
6051 GCACATTTC CCGAAAAGTG CCACCTGACG TCTAAGAAAC CATTATTATC
6101 ATGACATTAA CCTATAAAA TAGGCGTATC ACGAGGCCCT TTCGTCTTCA
6151 A

35/42

FIG. 13A

1 GAATTCTTAC ACTTAGTTAA ATTGCTAACT TTATAGATTA CAAAACCTAG
51 GAAATCGATT TGGATGAAAA AAGTAGTACT GGGCAAAAAA GGGGATACAG
101 TGGAACTGAC CTGTACAGCT TCCAGAAGA AGAGCATACA ATTCCACTGG
151 AAAA ACTCCA ACCAGATAAA GATTCTGGGA AATCAGGGCT CCTTCTTAAC
201 TAAAGGTCCA TCCAAGCTGA ATGATCGCGC TGA CTCAAGA AGAAGCTTGT
251 GGGACCAAGG AA ACTTTCCC CTGATCATCA AGAATCTTAA GATAGAAGAC
301 TCAGATACTT ACATCTGTGA AGTGGAGGAC CAGAAGGAGG AGGTGCAATT
351 GCTAGTG TTC GGATTGACTG CCAACTCTGA CACCCACCTG CTTCAGGGGC
401 AGAGCCTGAC CCTGACCTTG GAGAGCCCCC CTGGTAGTAG CCCCTCAGTG
451 CAATGTAGGA GTCCAAGGGG TAAAACATA CAGGGGGGA AGACCCCTCTC
501 CGTGTCTCAG CTGGAGCTCC AGGATAGTGG CACCTGGACA TGCACTGTCT
551 TGCAGAACCA GAAGAAGGTG GAGTTCAAAA TAGACATCGT GGTGCTAGCT
601 TTCCAGAAGG GGAAGATCTT TCCCGAGGGC GGCAGCCTGG CCGCGCTGAC
651 CGCGCACCAG GCTTGCCACC TGCCGCTGGA GACTTTCACC CGTCATCGCC

36/42

FIG. 13B

701 AGCCGCGCGG CTGGGAACAA CTGGAGCAGT GCGGCTATCC GGTGCAGCGG
751 CTGGTCGCCC TCTACCTGGC GCGCGGGCTG TCGTGGAACC AGGTCGACCA
801 GGTGATCCGC AACGCCCTGG CCAGCCCCCG CAGCGGCGGC GACCTGGGCG
851 AAGCGATCCG CGAGCAGCCG GAGCAGGCC GTCTGGCCCT GACCCCTGGCC
901 GCCGCCGAGA GCGAGCGCTT CGTCCGGCAG GGCACCGGA ACGACGAGGC
951 CGGCGCGGCC AACGCCGACG TGGTGAGCCT GACCTGCCCG GTCGCCGCCG
1001 GTGAATGCGC GGGCCCCGGC GACAGCGCG ACGCCCTGCT GGAGCGCAAC
1051 TATCCCACTG GCGCGGAGTT CCTCGCGGAC GCGGGCGACG TCAGCTTCAG
1101 CACCCGCGGC ACGCAGAACT GGACGGTGA GCGGCTGCTC CAGGCGCACC
1151 GCCAACTGGA GGAGCGCGC TATGTGTTTCG TCGGCTACCA CGGCACCTTC
1201 CTCGAAGCGG CGCAAAGCAT CGTCTTCGGC GGGGTGCGG CGCGCAGCCA
1251 GGACCTCGAC GCGATCTGGC GCGGTTTCTA TATCGCCGGC GATCCGGCGC
1301 TGGCCTACGG CTACGCCCAG GACCAGGAAC CCGACGCACG CGGCCGGATC
1351 CGCAACGGTG CCTGCTGCG GGTCTATGTG CCGCGCTCGA GCCTGCCGGG

37/42

FIG. 13C

1401 CTTCTACCGC ACCAGCCTGA CCCTGGCCGC GCCGGAGGCG GCGGGCGAGG
1451 TCGAACGGCT GATCGGCCAT CCGCTGCCGC TGGCCTTGA CGCCATCACC
1501 GGCCCCGAGG AGGAAGGCGG GCGCCTGGAG ACCATTCTCG GCTGGCCGCT
1551 GGCCGAGCGC ACCGTGGTGA TTCCCTCGGC GATCCCCACC GACCCGCGCA
1601 ACGTCGGCGG CGACCTCGAC CCGTCCAGCA TCCCCGACAA GGAACAGGCG
1651 ATCAGCGCCC TGCCGGACTA CGCCAGCCAG CCCGGCAAAC CGCCGCGCGA
1701 GGACCTGAAG TAACTGCCGC GACCGGCCGG CTCCCTTCGC AGGAGCCGCG
1751 CTTCTCGGGG CCTGGCCATA CATCAGGTTT TCCTGATGCC AGCCCAATCG
1801 AATATGAATT CTCATCGATT TCCATGGGAT CCTGCAGCCC AGCTTGGGA
1851 CCCTAGAGGT CCCCTTTTTT ATTTTGTAA TTGGGAGATC CAATTCTCAT
1901 GTTTGACAGC TTATCATCGA AGCTAGCTTT AATGCGGTAG TTTATCACAG
1951 TTAAATTGCT AACGCAGTCA GGCACCGTGT ATGAAATCTA ACAATGCGCT
2001 CATCGTCATC CTCGGCACCG TCACCCCTGGA TGCTGTAGGC ATAGGCTTGG
2051 TTATGCCGGT ACTGCCGGGC CTCTTGCGGG ATATCGTCCA TTCCGACAGC

38/42

FIG. 13D

2101 ATCGCCAGTC ACTATGGCGT GCTGCTAGCG CTATATGCGT TGATGCAATT
2151 TCTATGCGCA CCCGTTCTCG GAGCACTGTC CGACCGCTTT GGCCGCCGCC
2201 CAGTCCCTGCT CGCTTCGCTA CTTGGAGCCA CTATCGACTA CGCGATCATG
2251 GCGACCAACAC CCGTCCTGTG GATTCTCTAC GCCGGACGCA TCGTGGCCGG
2301 CATCACCCGC GCCACAGGTG CGGTTGCTGG CGCCTATATC GCCGACATCA
2351 CCGATGGGA AGATCGGGT CGCCACTCG GGCTCATGAG CGCTTGTTTC
2401 GGCGTGGTA TGGTGCAGG CCCCCTGGCC GGGGACTGT TGGCGGCCAT
2451 CTCCTTGAC GCACCATTC TTGCGGCGC GGTGCTCAAC GGCTCAACC
2501 TACTACTGG CTGCTTCCTA ATGCAGGAGT CGCATAAGG AGAGCGTCGT
2551 CCGATGCCCT TGAGAGCCTT CAACCCAGTC AGCTCCTTCC GGTGGCGCG
2601 GGGCATGACT ATCGTCGCCG CACTTATGAC TGTCTTCTTT ATCATGCAAC
2651 TCGTAGGACA GGTGCCGGCA GCGCTCTGG TCATTTTCGG CGAGGACCGC
2701 TTTCGCTGGA GCGCGACGAT GATCGGCCCTG TCGCTTGCGG TATTCGGAAT
2751 CTTGCACGCC CTCGCTCAAG CCTTCGTCAC TGGTCCCCGCC ACCAAACGTT

39/42

FIG. 13E

2801 TCGGCGAGAA GCAGGCCATT ATCGCCGGCA TGGGGGCCGA CGGCTGGGC
2851 TACGTCTTGC TGGCGTTCGC GACGCGAGGC TGGATGCGCT TCCCCATTAT
2901 GATTCTTCTC GCTTCCGGCG GCATCGGGAT GCCCGCGTTG CAGGCCATGC
2951 TGTCCAGGCA GGTAGATGAC GACCATCAGG GACAGCTTCA AGGATCGCTC
3001 GCGGCTCTTA CCAGCCTAAC TTCGATCACT GGACCGCTGA TCGTCACGGC
3051 GATTATGCC GCCTCGGCGA GCACATGGAA CGGGTTGGCA TGGATTGTAG
3101 GCGCCGCCCT ATACCTTGTC TGCCCTCCCCG CGTTGCGTCG CGGTGCATGG
3151 AGCCGGGCCA CCTCGACCTG AATGGAAGCC GCGGGCACCT CGCTAACGGA
3201 TTCACCACTC CAAGAATTGG AGCCAATCAA TTCTTGCGGA GAACTGTGAA
3251 TCGGCAAACC AACCCTTGGC AGAACATATC CATCGCGTCC GCCATCTCCA
3301 GCAGCCGCAC GCGGCGCATC TCGGGGGATG ATCAGCTGCC TCGCGCGTTT
3351 CGGTGATGAC GGTGAAAACC TCTGACACAT GCAGCTCCCC GAGACGGTCA
3401 CAGCTTGCT GTAGCGGAT GCCGGGAGCA GACAAGCCCC TCAGGGCGCG
3451 TCAGCGGGTG TTGGCGGGTG TCGGGGGCGA GCCATGACCC AGTCACGTAG

FIG. 13F

3501 CGATAGCGGA GTGTATACTG GCTTAACTAT GCGGCATCAG AGCAGATTGT
3551 ACTGAGAGTG CACCATATGC GGTGTGAAAT ACCGCACAGA TGCGTAAGGA
3601 GAAAATACCG CATCAGGCGC TCTTCCGCTT CCTCGCTCAC TGA CT CGCTG
3651 CGCTCGGTG TTCGGCTGCG GCGAGCGGTA TCAGCTCACT CAAAGGCGGT
3701 AATACGGTTA TCCACAGAAAT CAGGGGATAA CGCAGGAAAG AACATGTGAG
3751 CAAAAGGCCA GCAAAGGCC AGGAACCGTA AAAAGGCCGC GTTGCTGGCG
3801 TTTTTCCTA GGCTCCGCCC CCTGACGAG CATCACAAA ATCGACGCTC
3851 AAGTCAGAGG TGGCGAAACC CGACAGGACT ATAAAGATAC CAGGCGTTTC
3901 CCCCTGGAAG CTCCCTCGTG CGCTCTCCTG TTCCGACCCT GCCGCTTACC
3951 GGATACCTGT CCGCCCTTCT CCCTTCGGGA AGCGTGCGC TTTCTCAATG
4001 CTCACGCTGT AGGTATCTCA GTTCGGTGTA GGTGCTTCGC TCCAAGCTGG
4051 GCTGTGTGCA CGAACCCCCC GTTCAGCCCCG ACCGCTGCGC CTTATCCGGT
4101 AACTATCGTC TTGAGTCCAA CCCGGTAAGA CACGACTTAT CGCCACTGGC
4151 AGCAGCCACT GGTAACAGGA TTAGCAGAGC GAGGTATGTA GGCGGTGCTA
4201 CAGAGTTCTT GAAGTGGTGG CCTAACTACG GCTACACTAG AAGGACAGTA

40/42

41/42

FIG. 13G

4251 TTTGGTATCT GCGCTCTGCT GAAGCCAGTT ACCTTCGGAA AAAGAGTTGG
4301 TAGCTCTTGA TCCGGCAAAC AAACCACCGC TGGTAGCGGT GGTTTTTTTG
4351 TTTGCAAGCA GCAGATTACG CGCAGAAAAA AAGGATCTCA AGAAGATCCT
4401 TTGATCTTTT CTACGGGGTC TGACGCTCAG TGGAAACGAAA ACTCACGTTA
4451 AGGGATTTG GTCATGAGAT TATCAAAAAG GATCTTCACC TAGATCCTTT
4501 TCAGATCTCC CGATCTTTAG CTGCTCTGGT TTGCCCCAAG CGCATTTGCAT
4551 AATCTTTCAG GGTATGCGT TGTTCCTAC AACCTCCTTA GTACATGCAA
4601 CCATTATCAC CGCCAGAGGT AAAATAGTCA ACACGCACGG TGTTAGATAT
4651 TTATCCCCTG CGGTGATAGA TTTAACGTAT GAGCACAAAA AAGAAACCAT
4701 TAACACAAGA GCAGCTTGAG GACGCACGTC GCCTTAAAGC AATTATGAA
4751 AAAAAGAAAA ATGAACTTGG CTTATCCCAG GAATCTGTGC CAGACAAGAT
4801 GGGGATGGGG CAGTCAGGCG TTGGTGCTTT ATTTAATGGC ATCAATGCAT
4851 TAAATGCCTA TAACGCCGCA TTGCTTACAA AAATTCTCAA AGTTAGCGTT
4901 GAAGAATTTA GCCCTTCAAT CGCCAGAGAA ATCTACGAGA TGTATGAAGC
4951 GGTAGTATG CAGCCGTCAC TTAGAAGTGA GTATAGTAC CCTGTTTTTT

FIG. 13H

5001 CTCATGTTCA GGCAGGGATG TTCTCACCTA AGCTTAGAAC CTTTACCAAA
5051 GGTGATGCGG AGAGATGGGT AAGCACAACC AAAAAAGCCA GTGATTCTGC
5101 ATTCTGGCTT GAGGTTGAAG GTAATTCCAT GACCGCACCA ACAGGCTCCA
5151 AGCCAAGCTT TCCTGACGGA ATGTTAATTC TCGTTGACCC TGAGCAGGCT
5201 GTTGAGCCAG GTGATTCTG CATAGCCAGA CTTGGGGTG ATGAGTTTAC
5251 CTTCAAGAAA CTAATTAGGG ATAGCGGTCA GGTGTTTTTA CAACCACTAA
5301 ACCCACAGTA CCAATGATC CCATGCAATG AGAGTTGTTC CGTTGTGGGG
5351 AAAGTTATCG CTAGTCAGTG GCCTGAAGAG ACGTTTGGCT GATCGGCAAG
5401 GTGTTCTGGT CGGCGCATAG CTGATAACAA TTGAGCAAGA ATCTTCATCG
5451 GGGCTGCAGC CCACGATGCG TCCGGCGTAG AGGATCTCTC ACCTACCAAA
5501 CAATGCCCCC CTGCAAAAAA TAAATTCATA TAAAAACAT ACAGATAACC
5551 ATCTGCGGTG ATAAATTATC TCTGGCGGTG TTGACATAAA TACCACTGGC
5601 GGTGATACTG AGCACATCAG CAGGACGCAC TGACCACCAT GAAGGTGACG
5651 CTCCTAAAAAT TAAGCCCTGA AGAAGGGCAG CATTCAAAGC AGAAGGCTTT
5701 GGGGTGTGTG ATACGAAACG AAGCATT

INTERNATIONAL SEARCH REPORT

International Application No. PCT/US91/02954

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all) ⁶

According to International Patent Classification (IPC) or to both National Classification and IPC

IPC(5): C07H 15,/12; C12N 1/00; C12P 21/06; C07K 3/00; A61K 37/22; C12Q 1/00
U.S.CL.: 536/27; 435/243; 530/350; 435/69.1

II. FIELDS SEARCHED

Minimum Documentation Searched ⁷

Classification System	Classification Symbols
U.S.	536/27; 435/243; 530/350; 435/69.1

Documentation Searched other than Minimum Documentation
to the Extent that such Documents are Included in the Fields Searched ⁸

AUTOMATED PATENT FILES USPAT: COMMERCIAL DATABASES STN AND DIALOG

III. DOCUMENTS CONSIDERED TO BE RELEVANT ⁹

Category [*]	Citation of Document, ¹¹ with indication, where appropriate, of the relevant passages ¹²	Relevant to Claim No. ¹³
Y	AM. REV. RESPIR. DIS, VOL. 138 (2). ISSUED 1988, LIND ET AL pages 429-434 ABSTRACT ONLY: SEE ENTIRE DOCUMENT.	1-55
T	US.A. 4,999,344 (JETT-TILTON ET AL.) 12 MARCH 1991, SEE ENTIRE DOCUMENT.	1-55
Y	THE JOURNAL OF BIOLOGICAL CHEMISTRY, VOL. 263, NO. 2, ISSUED 15 JANUARY 1988. ANDRE' ET AL. "SEVERIN, GELSOIN, AND VILLIN SHARE A HOMOLOGOUS SEQUENCE IN REGIONS PRESUMED TO CONTAIN F-ACTIN SERVING DOMAINS". PP 722-727: SEE ENTIRE DOCUMENT.	1-55
Y	METHODS IN ENZYMOLOGY, VOL. 25, ISSUED 1972, M.H. KLAPPE ET AL. "ACYLATION WITH DICARBOXYLIC ACID ANHYDRIDES". PP. 571-60: SEE ENTIRE DOCUMENT.	1-55

* Special categories of cited documents: ¹⁰

"A" document defining the general state of the art which is not
considered to be of particular relevance

"E" earlier document but published on or after the international
filing date

"L" document which may throw doubts on priority claim(s) or
which is cited to establish the publication date of another
citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or
other means

"P" document published prior to the international filing date but
later than the priority date claimed

"T" later document published after the international filing date
or priority date and not in conflict with the application but
cited to understand the principle or theory underlying the
invention

"X" document of particular relevance: the claimed invention
cannot be considered novel or cannot be considered to
involve an inventive step

"Y" document of particular relevance: the claimed invention
cannot be considered to involve an inventive step when the
document is combined with one or more other such docu-
ments, such combination being obvious to a person skilled
in the art.

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search

12 July 1991

Date of Mailing of this International Search Report

06 SEP 1991

International Searching Authority

ISA/US

Signature of Authorized Officer

Bradley L. Sisson
Bradley L. Sisson

ANNEX M3

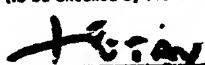
International Application No: PCT/US91 / 02954

MICROORGANISMS	
Optional Sheet in connection with the microorganism referred to on page <u>49</u> , line <u>26</u> of the description ¹	
A. IDENTIFICATION OF DEPOSIT: <u>pCD4-gelsolin</u>	
Further deposits are identified on an additional sheet <input type="checkbox"/> ²	
Name of depositary institution ⁴ <u>In Vitro International, Inc.</u>	
Address of depositary institution (including postal code and country) ⁴ <u>611 P. Hammonds Ferry Road</u> <u>Linthicum, Maryland 21090 (USA)</u>	
Date of deposit ⁴ <u>May 4, 1990</u>	Accession Number ⁴ <u>IVI-10253</u>
B. ADDITIONAL INDICATIONS ⁷ (leave blank if not applicable). This information is continued on a separate attached sheet <input type="checkbox"/>	
<p>In respect of those designations in which a European patent is sought samples of the deposited microorganisms will be made available until the publication of the mention of the grant of the European patent or until the date on which the application has been refused or withdrawn or is deemed to be withdrawn only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28(4) EPC).</p>	
C. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE ⁸ (if the indications are not for all designated States)	
<u>EPO</u>	
D. SEPARATE FURNISHING OF INDICATIONS ⁹ (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later ⁹ (Specify the general nature of the indications e.g., "Accession Number of Deposit")	
<p>E. <input type="checkbox"/> This sheet was received with the international application when filed (to be checked by the receiving Office)</p> <p style="text-align: center;"><u>HEITAO</u> (Authorized Officer)</p> <p><input type="checkbox"/> The date of receipt (from the applicant) by the International Bureau is <u>19 AUG 1991</u></p> <p style="text-align: center;">was (Authorized Officer)</p>	

January 1985)

ANNEX M3

International Application No: PCT/ US91 / 02954

MICROORGANISMS	
Optional Sheet in connection with the microorganism referred to on page <u>49</u> , line <u>27</u> of the description ¹	
A. IDENTIFICATION OF DEPOSIT ¹ <u>pl70.2</u> Further deposits are identified on an additional sheet <input type="checkbox"/> ²	
Name of depositary institution ⁴ <div style="text-align: center; padding: 10px;">In Vitro International, Inc.</div>	
Address of depositary institution (including postal code and country) ⁴ <div style="text-align: center; padding: 10px;">611 P. Hammonds Ferry Road Linthicum, Maryland 21090 (USA)</div>	
Date of deposit ⁵ May 4, 1990	Accession Number ⁶ IVI-10252
B. ADDITIONAL INDICATIONS ⁷ (leave blank if not applicable). This information is continued on a separate attached sheet <input type="checkbox"/>	
<p>In respect of those designations in which a European patent is sought samples of the deposited microorganisms will be made available until the publication of the mention of the grant of the European patent or until the date on which the application has been refused or withdrawn or is deemed to be withdrawn only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28(4) EPC).</p>	
C. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE ⁸ (if the indications are not for all designated States)	
EPO	
D. SEPARATE FURNISHING OF INDICATIONS ⁹ (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later ⁹ (Specify the general nature of the indications e.g. "Accession Number of Deposit")	
E. <input type="checkbox"/> This sheet was received with the international application when filed (to be checked by the receiving Office)	
<div style="margin-bottom: 10px;">  (Authorized Officer) </div> <div style="margin-bottom: 10px;"> <input type="checkbox"/> The date of receipt (from the applicant) by the International Bureau is 19 AUG 1991 </div> <div style="margin-bottom: 10px;"> was _____ (Authorized Officer) </div>	

January 1985)